Functional Characterisation of the Genes Mutated in Dyskeratosis Congenita

Richard William Beswick, BSc (Hons)

Thesis submitted to the University of London for the degree of Doctor of Philosophy

February 2013

Centre for Paediatrics
Blizard Institute
Barts and The London School Medicine and Dentistry
Queen Mary University of London
London
UK

In loving memory of my mum

Lynda Beswick

(A promise fulfilled)

Abstract

Dyskeratosis congenita (DC) is a multi system disorder that exhibits considerable clinical and genetic heterogeneity. It is characterised by mucocutaneous features, bone marrow failure and a predisposition to cancer. Research has identified mutations affecting several telomerase components and patients often have short telomeres, implicating defective telomere maintenance in this disease. Affected components include dyskerin, NOP10 and NHP2, which together with GAR1 form a protein core common to telomerase and all other H/ACA ribonucleoprotein complexes (H/ACA RNPs). Initially characterised as H/ACA RNP components important for pseudouridylation and rRNA processing, their role in the functionally distinct telomerase complex and telomere maintenance is less defined.

In order to better understand their implications in DC, this study investigated the importance of these core proteins for the integrity and function of telomerase in human cells. RNAi knockdown studies demonstrated that dyskerin, NOP10 and NHP2 are necessary for the accumulation of TERC (telomerase RNA component); dyskerin and NOP10 for telomerase activity. Moreover, dyskerin was found to be important for maintaining telomere length over time. The impact of *NOP10* and *NHP2* missense mutations was also analysed *in vitro*, which indicated that they impair TERC accumulation. The potential effect on pseudouridylation was also considered in this study; the analysis of other H/ACA RNA levels in these knockdown experiments and in a cohort of patients with *DKC1* mutations revealed an irregular and inconsistent impact compared to that observed on TERC.

Finally, defective telomere maintenance is heavily implicated as the primary cause of DC and very short telomeres have been proposed as a diagnostic marker. This study investigated telomere length in a patient cohort of unprecedented size. It demonstrated the prevalence of the telomere length defect, but telomere length was not found to correlate with either genetic subtype or disease severity, implicating the rate of telomere shortening as the correlating factor instead.

Acknowledgements

I would first like to thank my supervisors Prof. Inderjeet Dokal and Dr Tom Vulliamy for the opportunity to carry out this work which was funded by the Wellcome Trust. I also thank them for their support and guidance over the years but especially for their encouragement. It was much needed at times.

I give special thanks to my good friend Dr Mike Kirwan, not only for his invaluable help and support on many aspects of this project but for so much more. I am indebted to him.

I am very grateful to Dr Amanda Walne and Dr Anna Marrone for their advice and contributions, not to mention their friendship and support. Also, all those in the centre for paediatrics and in the haematology laboratories for their help when it was needed, but most of all for making me feel at home amongst them. It was a pleasure to work with them all.

Finally, a very big thank you to all my family, especially my parents Lynda and John, for their continued supported and encouragement ‘come rain or shine’. I owe them so much.

Publications arising from this Work

Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. Walne, A. J., T. Vulliamy, A. Marrone, **R. Beswick**, M. Kirwan, Y. Masunari, F.-h. Al-Qurashi, M. Aljurf and I. Dokal (2007). *Hum Mol Genet* **16**(13): 1619-1629.

Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Vulliamy, T. J., ****R. Beswick**, M. Kirwan, A. Marrone, M. Digweed, A. Walne and I. Dokal (2008). *Proc Natl Acad Sci USA* **105**(23): 8073-8.
*Joint first author

Differences in disease severity but similar telomere lengths in genetic subgroups of patients with telomerase and shelterin mutations. Vulliamy, T. J., M. J. Kirwan, **R. Beswick**, U. Hossain, C. Baqai, A. Ratcliffe, J. Marsh, A. Walne and I. Dokal (2011). *PloS one* **6**(9): e24383.

Telomere length measurement can distinguish pathogenic from non-pathogenic variants in the shelterin component, TIN2. Vulliamy, T., **R. Beswick**, M. J. Kirwan, U. Hossain, A. J. Walne and I. Dokal (2011). *Clin Genet* **81**(1): 76-81.

Mutations in *C16orf57* and normal-length telomeres unify a subset of patients with dyskeratosis congenita, poikiloderma with neutropenia and Rothmund-Thomson syndrome. Walne, A. J., T. Vulliamy, **R. Beswick**, M. Kirwan and I. Dokal (2010). *Hum Mol Genet* **19**(22): 4453-61.

List of Abbreviations

ALT	alternative lengthening of telomeres
AA	Aplastic Anaemia
AD	autosomal dominant
AR	autosomal recessive
bp	base pair(s)
BM	bone marrow
BMF	bone marrow failure
BSA	bovine serum albumin
CSC	chondroitin sulphate C
cDNA	complementary DNA
C_T	cycle threshold
dCTP	deoxycytodine triphoshpate
DNA	deoxyribonucleic acid
dNTP	deoxyribonucletoide triphoshpate
DMSO	dimethyl sulphoxide
DTT	dithiothreitol
dsDNA	double stranded DNA
dsRNA	double stranded RNA
DMEM	Dulbecco's modified eagle medium
DC	dyskeratosis congenita
DCR	dyskeratosis congenita registry
ECL	enhanced chemiluminescence
EDTA	ethylene diamine tetra-acetic acid
FITC	fluorescein isothiocyanate
FISH	fluorescence in situ hybridisation
FCS	foetal calf serum
FSC	forward scatter
g	acceleration of gravity
g	gram
GFP	green fluorescent protein
HRP	horseradish peroxidise

hr	hour(s)
HH	Hoyeraal Hreidarsson syndrome
kb	kilobase pair
kDa	kilodalton
l	litre
LB	Luria-Bertani broth
mRNA	messenger RNA
mRNA	messenger RNA
miRNA	micro RNA
µg	microgram
µl	microlitre
µM	micromolar
mg	milligram
ml	millilitre
min	minute(s)
M	molar
M-MLV RT	moloney murine leukaemia virus reverse transcriptase
MMqPCR	monochrome multiplex PCR
mTR	mouse telomerase RNA
MDS	Myelodysplastic syndrome (myelodysplasia)
ng	nanogram
nM	nanomolar
NTC	non-targeting control
PNA	peptide nucleic acid
PMSF	phenylmethylsulphonyl fluoride
PBS	phosphate buffered saline
PN	poikiloderma with neutropenia
PAGE	polyacrylamide gel electrophoresis
PB	polybrene
PCR	polymerase chain reaction
PVP-40	Polyvinylpyrrolidone 40
Pf	*Pyrococcus furiosus*
qPCR	quantitative real time PCR

RIPA	Radio-ImmunoPrecipitation assay
RCLB	red cell lysis buffer
RTL	relative telomere length
Rn	reporter normalised
RE	restriction enzyme
RT	reverse transcriptase
rpm	revolutions per minute
RNA	ribonucleic acid
RNP	ribonucleoprotein complex
rRNA	ribosomal RNA
RNAi	RNA interference
RTS	Rothmund-Thomson syndrome
SSC (Southern blot)	saline sodium citrate
sec	second(s)
shRNA	short hairpin RNA
SSC (flow FISH)	side scatter
ssDNA	single stranded DNA
scaRNA	small cajal body RNA
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
SD	Standard deviation
SDS	sodium dodecyl sulphate
TRAP	telomerase repeat amplification protocol
TERT	telomerase reverse transcriptase
TERC	telomerase RNA component
TEMED	tetramethylethylenediamine
TAE	tris acetate EDTA
TBS	tris buffered saline
TE	tris EDTA
TBE	tris-borate-EDTA
UV	ultraviolet
UTR	untranslated region
yrs	years

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Dyskeratosis Congenita

First described over 100 years ago dyskeratosis congenita (DC), also known as Zinsser-Engman-Cole syndrome, is a rare inherited multi system disorder classically characterised by a triad of mucocutaneous features: nail dystrophy, leucoplakia and abnormal skin pigmentation (Drachtman and Alter 1995). Considered to be a rare bone marrow failure (BMF) syndrome due to the prevalence of bone marrow (BM) abnormalities, the incidence of age related features including premature hair loss/greying, predispositions to malignancy, pulmonary and skeletal abnormalities also allow DC to be considered a premature ageing syndrome (Dokal 2001). Genetic and molecular studies have subsequently indicated the primary underlying pathology of this disease to be a defect in telomere maintenance (reviewed Walne et al. 2005).

1.2. Clinical features of DC

DC classically presents with a mucocutaneous triad of abnormalities: abnormal skin pigmentation, nail dystrophy, and leucoplakia (figure 1.1), which form the basis for its diagnosis. However, a range of abnormalities in the dental, gastrointestinal, genitourinary, neurological, ophthalmic, pulmonary and skeletal systems are also reported in this multi system disorder (table 1). Progressing with age the onset of disease typically occurs during childhood, the nail and skin abnormalities often visible by 10yrs of age followed by leucoplakia in the early teens. During adolescence and early adulthood the more severe abnormalities begin to develop which vary considerably between patients, though 80-89% suffer BMF by 30yrs of age. Early mortality is also characteristic of DC, usually the result of BMF but also the pulmonary complications and malignancy associated with this disease; typically head, neck and gynaecological squamous cell carcinomas (particularly oral) as well as acute leukaemias and MDS (Drachtman and Alter 1995; Knight et al. 1998; Dokal 2000; Alter et al. 2009). Disease severity varies and is generally gauged by the age of BMF onset and the number of somatic abnormalities that manifest: BMF before the age of 10yrs and several other features of DC classed as severe; BMF after the age of 20yrs classed as mild (Dokal 2000). What's more, the diversity in clinical presentation and severity can create difficulty in diagnosing DC, an issue compounded by substantial clinical overlap with other syndromes (table 1) (reviewed Marrone and Dokal 2006).

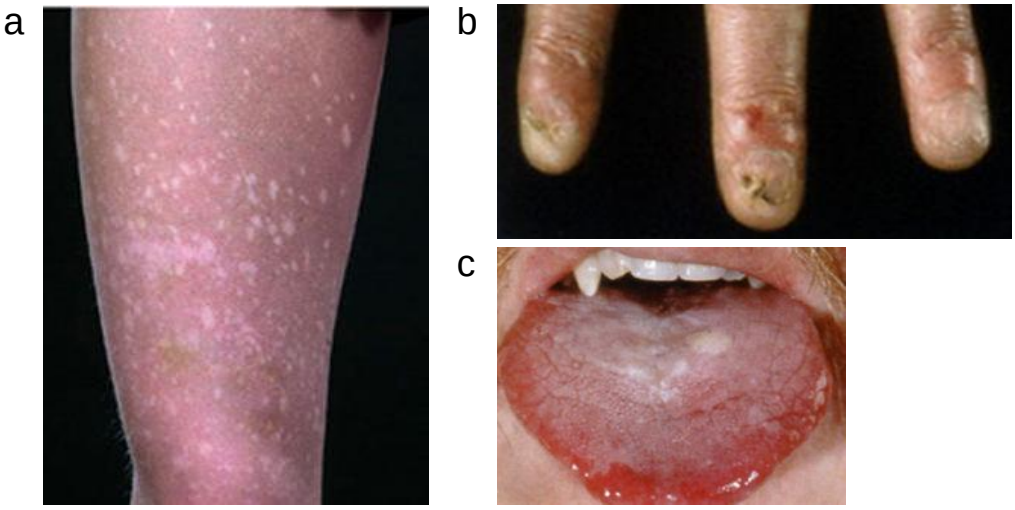


Figure 1.1. The mucocutaneous triad of DC, (a) abnormal skin pigmentation, (b) nail dystrophy and (c) leucoplakia.

Table 1. Clinical features of DC and their overlap with other syndromes.

Clinical features	BMF syndromes								Premature ageing syndromes				
	DC	AA	AML	MDS	PNH	DBA	FA	SDS	BS	CS	HGPS	RTS	WS
Characteristic													
Abnormal skin pigmentation	89%						✓		✓	✓	✓	✓	✓
Nail dystrophy	88%											✓	✓
Leucoplakia	78%						✓						
BMF	86%	✓	✓	✓	✓	✓	✓	✓					
Common													
Epiphora	31%												
Mental retardation	25%						✓						
Pulmonary disease	20%						✓		✓				
Short stature	20%					✓	✓	✓	✓	✓	✓	✓	✓
Dental carries/loss	17%							✓	✓			✓	✓
Oesophageal stricture	17%												
Premature hair loss/ greying	16%									✓		✓	✓
Hyperhidrosis	15%												
Less frequent													
Malignancy	10%					✓	✓	✓	✓			✓	✓
Liver disease	7%						✓	✓					
Ataxia/cerebellar hypoplasia	7%												
Hypogonadism	6%						✓		✓			✓	✓
Microcephaly	6%						✓			✓			
Urethral stricture/ phimosis	5%												
Osteoporosis	5%						✓	✓			✓		✓
Deafness	1%												

AA: Aplastic anaemia; AML: Acute myeloid leukaemia; PNH: Paroxysmal nocturnal haemoglobinuria; DBA: Diamond-Blackfan anaemia; FA: Fanconi anaemia; SDS: Shwachman-Diamond syndrome; BS: Bloom syndrome; CS: Cockayne syndrome; HGPS: Hutchinson-Gilford progeria syndrome; RTS: Rothmund-Thomson syndrome; WS: Werner syndrome. (from Marrone and Dokal 2006)

1.3. Genetic features of DC

DC exhibits three modes of inheritance X-linked recessive, autosomal dominant (AD) and autosomal recessive (AR). With the majority of cases reported in males, X-linked is the most common and is associated with the more classical severe form of the disease. AD and AR forms are more variable with AD being the rarest and mildest form of DC while AR can be severe in some cases. Clearly then DC has multiple genetic basis', each of which can be associated with a different degree of pathological impact and disease severity (Dokal 2000; Marrone and Mason 2003).

The Dyskeratosis Congenita Registry (DCR) was established at the Hammersmith Hospital in 1995 (relocated to Barts and The London in 2006) to facilitate research into the underlying genetics and mechanism of DC disease. At the point of commencing this study five different genes had been implicated in DC, two of which were unpublished at the time. These are the X-linked recessive gene *DKC1* (Heiss et al. 1998), the AD genes *TERC* and *TERT* (Vulliamy et al. 2001b; Armanios et al. 2005; Vulliamy et al. 2005) and the unpublished AR genes *NOP10* and *NHP2*. However, collectively these genes are not sufficient to explain all cases of the disease. X-linked DC represents ~32% of DCR patients and AD-DC only ~5% with the remaining 63% a mixture of AR inheritance and sporadic cases (Marrone and Dokal 2006), of which the *NOP10* and *NHP2* mutations account for only one and two families respectively. Therefore, the genetic basis for a large proportion of the DCR remains elusive.

DKC1, *NOP10* and *NHP2* encode the highly conserved nucleolar proteins dyskerin, NOP10 and NHP2. These are all components of H/ACA ribonucleoprotein complexes (H/ACA RNPs) which are important for ribosomal RNA (rRNA) processing and pre messenger RNA (mRNA) splicing (Meier 2005). Similarly, these proteins are also core components of the telomerase complex, a specialised H/ACA RNP serving a uniquely different role in telomere maintenance. Through telomerase this functional link can be extended to include *TERC* and *TERT* which encode the complexes distinguishing components, the telomerase RNA component (TERC) and telomerase reverse transcriptase (TERT) (Collins and Mitchell 2002). Together, these five genes implicate two structurally related yet functionally distinct complexes in the molecular pathology of DC, but particularly telomerase and telomere maintenance.

1.4. H/ACA ribonucleoprotein complexes

H/ACA ribonucleoprotein complexes (H/ACA RNPs) are a class of small RNP complexes involved in the post translation modification of RNAs, specifically the isomerisation of specific uridine residues to pseudouridines in a process known as pseudouridylation. The exact purpose of this modification is unclear but is thought to increase stability of the target RNA and a break down in this process has been shown to severely impair rRNA processing and ribosomal assembly. H/ACA RNPs are characterised by their H/ACA RNA component which guides this process through site directed base pairing with the target RNA. In this way the many different species of H/ACA RNA creates diversity among the H/ACA RNPs and allow them to collectively target a multitude of different uridine residues. The H/ACA RNAs themselves are defined by specific H (ANANNA) and ACA box sequence motifs and form secondary structures that feature a double hairpin containing pseudouridylation pockets in which the target specific sequence elements are located (figure 1.2).

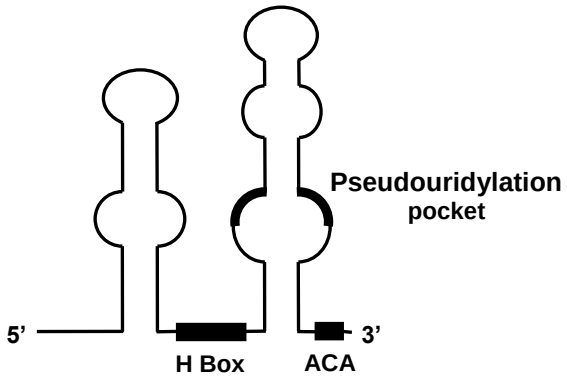


Figure 1.2. A model of H/ACA RNA secondary structure.

Most H/ACA RNPs localise to the nucleoli where they target rRNAs and are referred to as small nucleolar RNPs (H/ACA snoRNPs), their RNAs as H/ACA snoRNAs. However, some H/ACA RNAs contain a CAB box sequence motif which targets the RNP for localisation to Cajal bodies where they target spliceosomal small nuclear RNAs (snRNAs) and are referred to as small Cajal body RNPs (H/ACA scaRNPs), their RNAs as H/ACA scaRNAs. Despite differences in their localisation and RNA component all H/ACA RNPs comprise the same four proteins at their core: dyskerin, NOP10, NHP2, and GAR1, which are evolutionarily conserved with orthologues in

many species including mice, yeast and archaea. (reviewed Filipowicz and Pogacic 2002; Henras et al. 2004a; Meier 2005).

Initially characterised as H/ACA RNP components in yeast, through the orthologues: Cbf5p (dyskerin), Nop10p (NOP10), Nhp2p (NHP2), and Gar1p (GAR1), these proteins were all found to be essential for pseudouridylation and rRNA processing. Sharing significant homology with the bacterial pseudouridine synthase TruB, Cbf5p was identified as the enzymatic component of the complex but was also shown to be important for RNP stability and H/ACA RNA accumulation along with Nop10p and Nhp2p (Girard et al. 1992; Bousquet-Antonelli et al. 1997; Henras et al. 1998; Lafontaine et al. 1998; Watkins et al. 1998; Dez et al. 2001; Henras et al. 2004b). Since then research in mice has revealed much about the interaction of these H/ACA RNP components. Assembly of the H/ACA RNP is indicated to takes place at the site of H/ACA RNA transcription in the nucleus. The protein NAF1 is implicated in chaperoning NAP57 (dyskerin) to this location where it associates with NOP10 and NHP2 (via NOP10) to form a core protein trimer that recruits the newly transcribed H/ACA RNA. GAR1 does not seem to play a part at this stage, associating with the complex post RNA recruitment to create a mature, functional H/ACA RNP (figure 1.3). At this point NAF1 is no longer associated with the complex and the discovery that GAR1 and NAF1 competitively bind to NAP57 suggests that these two proteins interchange during this maturation process (Wang and Meier 2004; Darzacq et al. 2006). Indeed, this may serve to explain why GAR1 was not found to be important for H/ACA RNA stability and accumulation.

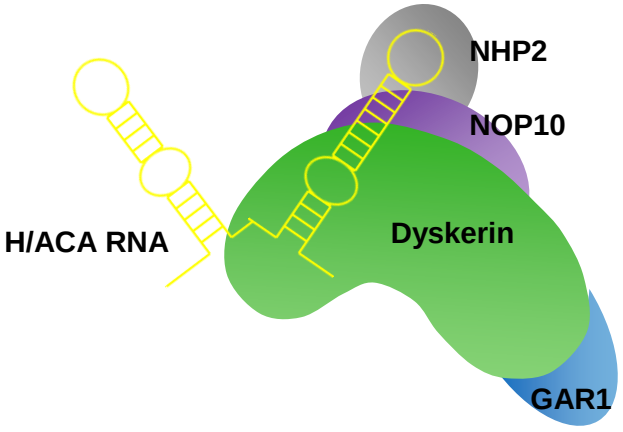


Figure 1.3. A schematic representation of an H/ACA RNP.

Taken together the data suggests a system in which the protein components associate first before recruiting a newly synthesised H/ACA RNA and that RNP integration is necessary for the H/ACA RNAs stability and accumulation. Conversely, studies have shown that NAP57, NOP10, NHP2 and GAR1 are capable of forming a protein only complex that does not require an associated H/ACA RNA for its stability (Wang and Meier 2004). Further, H/ACA RNAs have been shown to bind irreversibly to the RNP, specifically NAP57, indicating that once formed H/ACA RNP's do not exchange their RNA components (Wang and Meier 2004; Kittur et al. 2006).

1.5. Telomeres and the telomerase complex

Telomeres are DNA-protein complexes that cap the ends of eukaryotic chromosomes and their function is essentially to protect the stability of the genome by preserving the integrity of the chromosome. The DNA component is a long stretch of 6bp TTAGGG repeats that is processed by unidentified nuclease activity to terminate in a single stranded 3' overhang. This telomeric DNA is folded back on itself, the ssDNA overhang invading the dsDNA to form a closed loop structure at the end of the chromosome called the T-loop. The exact mechanism behind this loop formation is not fully understood but is thought to be mediated by shelterin, a protein complex that associates along the length of the telomere (figure 1.4) (reviewed Blackburn 2001; de Lange 2005; Shay and Wright 2005). This effectively provides a 'neat and tidy' end to the chromosome that protects it from random degradation and fusion events, disguising it from DNA repair mechanisms that might otherwise mistake it for a chromosomal break point.

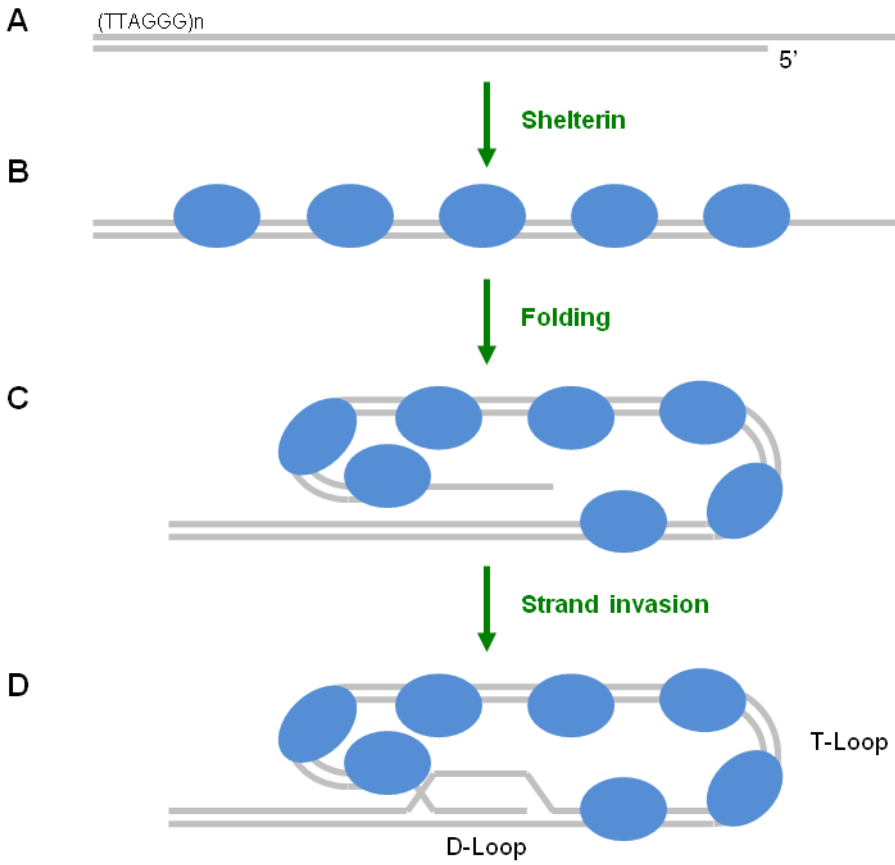


Figure 1.4. T-Loop formation and the proposed role of shelterin. (A) The telomere is extended during DNA replication and processed to create a long 3' overhang. (B) Shelterin proteins associate to form complexes along the length of the telomere. (C) Thought to be mediated by shelterin, the telomere is folded back on itself forming the T-loop. (D) The T-loop is closed by the 3' overhang invading the dsDNA of the telomere, creating a smaller D-loop in the process.

Telomeres also protect against the degradation that occurs with DNA replication. Semi-conservative DNA replication, or more specifically its 5'-3' nature, gives rise to the 'end replication problem' where the 3' lagging strand is unable to fully replicate causing 50-100bp to be lost from the end of the DNA with each division (figure 1.5). Telomeric DNA provides a 'buffer zone' protecting the genome from this slow attrition (reviewed Blackburn 2001; Shay and Wright 2005). The consequence of this is progressive telomere shortening as cells divide which can be observed with ageing both *in vivo* and *in vitro* (Harley et al. 1990; Vaziri et al. 1994; Allsopp et al. 1995). When telomeres are reduced to a critical length further cell divisions are prevented

and the cell enters replicative senescence or apoptosis (reviewed Blackburn 2001; Shay and Wright 2005) This correlates with earlier observations that normal somatic cells undergo a finite number of divisions the maximum possible called the Hayflick limit (Hayflick 1965). Telomere shortening is a particular threat to stem cell pools and proliferating tissues where cells replicate frequently, haematopoietic stem cells in the bone marrow being a relevant example. The answer to this problem for many of these cell types is telomerase.

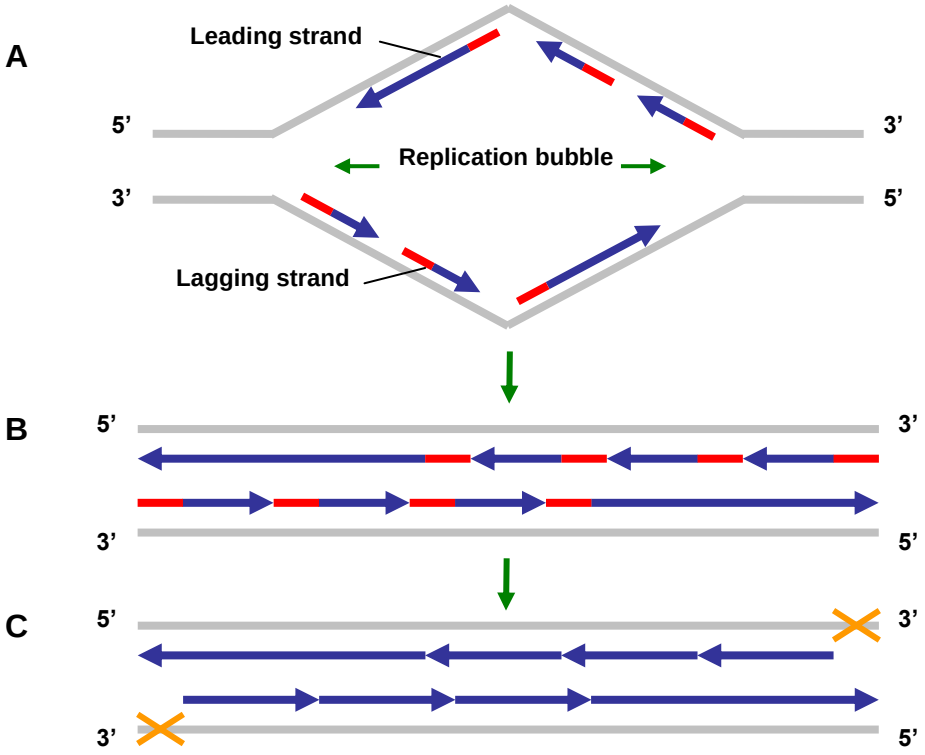


Figure 1.5. The end replication problem. (A) A replication bubble forms and DNA is replicated outward in both directions. (B) Synthesis of the lagging strand is fragmented - Okasaki fragments. (C) Primers (red) are removed and the gaps filled. This is not possible at the 5' end of the daughter strand leading to incomplete replication of the parental 3' strand and progressive DNA loss.

Telomerase is an RNP complex with a specialised reverse transcriptase (RT) function that acts to extend the telomere repeat region during DNA replication, allowing telomere length to be maintained through successive cell divisions and protecting the cell from telomere related senescence or apoptosis. It is composed of several subunits

but two in particular are responsible for its functional activity: a catalytic component TERT provides the enzymatic RT activity; and an RNA component TERC which provides the template for telomere elongation. Telomerase was also found to comprise an H/ACA RNP at its core (the telomerase RNP) via discovery of the H/ACA motif towards the 3' end of TERC, subsequent studies confirming its interaction with dyskerin, NOP10, NHP2 and GAR1 like any other H/ACA RNA. The telomerase RNP then recruits TERT to create a mature, functional telomerase complex (figure 1.6) (Reviewed Collins and Mitchell 2002; Cong et al. 2002). However, the role of telomerase is uniquely different to that of other H/ACA RNPs and no pseudouridylation targets have been identified to imply that TERC has a dual role. Therefore, in the context of telomerase, the catalytic function of dyskerin appears to be redundant suggesting that in this instance the role of the H/ACA RNP is simply to provide a structural framework for the larger complex. Certainly, like other H/ACA RNAs, in both yeast and mice the stable accumulation of TERC has been shown to depend on dyskerin, NOP10 and NHP2 (Dez et al. 2001; Wang and Meier 2004).

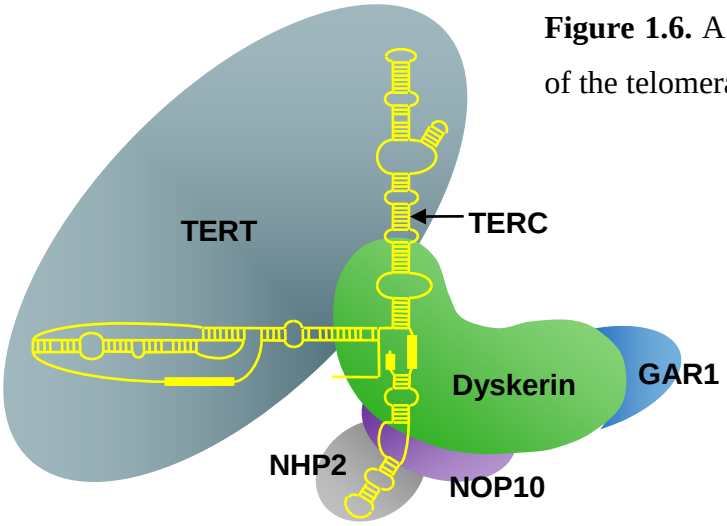


Figure 1.6. A schematic representation of the telomerase complex.

1.6. DC: a defect of telomere maintenance?

The first DC mutations characterised were in the X-linked dyskerin gene *DKC1*. A highly conserved protein, dyskerin was identified as the human orthologue to Cbf5p in yeast and the key catalytic component of human H/ACA RNPs (Heiss et al. 1998; Dragon et al. 2000; Pogacic et al. 2000). Structurally it has two key functional

domains: a TruB domain responsible for its catalytic activity; and a PUA domain that is important for RNA binding (Heiss et al. 1998; Aravind and Koonin 1999). Most of the dyskerin mutations identified in DC cluster in or around the PUA domain suggesting that they impact on dyskerin-H/ACA RNA interactions (Rashid et al. 2006; Vulliamy et al. 2006). Initially the association with telomerase was unknown and DC was thought to stem from a defect in pseudouridylation and rRNA processing (Luzzatto and Karadimitris 1998). However, the discovery that dyskerin is also part of the telomerase complex offered an alternative possibility and gave rise to renewed thought on the cause of DC. Indeed, subsequent research on cells from patients with *DKC1* mutations revealed no discernible differences in H/ACA RNA accumulation, pseudouridine content or rRNA processing between affected and unaffected family members. Instead patients exhibit a marked decrease in TERC accumulation, telomerase activity and telomere length (Mitchell et al. 1999b; Vulliamy et al. 2001a; Montanaro et al. 2002; Wong et al. 2004; Wong and Collins 2006).

Following this, heterozygous mutations in the AD form of DC were discovered in a second gene *TERC*, encoding the telomerase RNA: TERC (Vulliamy et al. 2001b). A functional RNA molecule, TERC undergoes rapid 3' end processing from the nascent RNA transcript to the mature functional molecule with a complex secondary structure. Within this structure are four key elements of functional significance (figure 1.7). Two are important for the catalytic activity of telomerase, the pseudoknot and the CR4-CR5 domains. The other two appear to be important for correct processing and accumulation of the mature TERC, these are the H/ACA and CR7 domains (reviewed Chen and Greider 2004). Mutations have been found in all of these elements and also deletions in both the 5' and 3' regions, though most locate to the pseudoknot domain (Marrone and Dokal 2006). As there are no pseudouridylation targets associated with TERC, pseudouridylation defects are not implicated by these mutations. Instead these patients also exhibit shorter than average telomeres and there is functional data to demonstrate that TERC mutations reduce telomerase activity through diminished TERC stability and accumulation or direct disruption to the catalytic process (Vulliamy et al. 2001b; Marrone et al. 2004; Vulliamy et al. 2004). In the heterozygous state these mutations do not have a dominant negative effect but instead cause haploinsufficiency, the unaffected allele proving inadequate for maintaining optimal telomerase activity. The result is gradual shortening of the telomeres that

gives rise to a delayed clinical presentation and disease anticipation in affected families, where the manifestation and severity of the disease increases with successive generations (Marrone et al. 2004; Vulliamy et al. 2004).

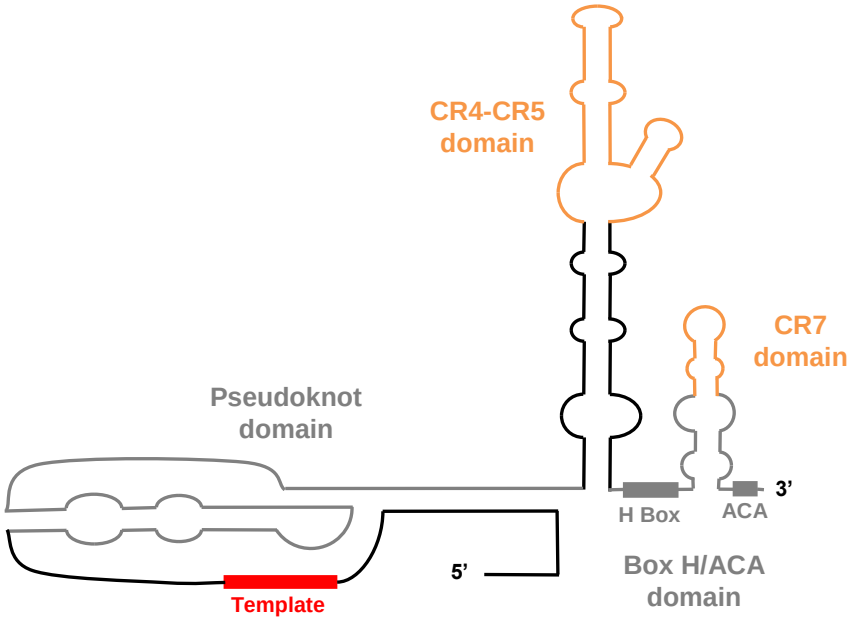


Figure 1.7. A model of TERC's secondary structure highlighting key domains

Heterozygous mutations have also been found in the *TERT* gene of some AD DC patients (Armanios et al. 2005; Vulliamy et al. 2005). TERT is a large protein with several functionally conserved regions. The characteristic, highly conserved RT motif and an adjacent telomerase specific T-motif are located in the C-terminal half of the protein (Reviewed Cong et al. 2002). Heterozygous mutations in TERT have been associated with reduced telomerase activity by haploinsufficiency, but generally the disease significance of TERT mutations is variable and considered more a risk factor than a cause (Armanios et al. 2005; Vulliamy et al. 2005; Yamaguchi et al. 2005).

Finally, homozygous mutations in the AR DC genes *NOP10* and *NHP2* are very rare with instances in only one and two families respectively. The NOP10 protein measures only 64 amino acids in length and has no known functional motifs (Henras et al. 1998). NHP2 is an RNA binding protein, its specificity for H/ACA RNAs determined indirectly through its association with the H/ACA RNP (Wang and Meier 2004). Though other H/ACA RNAs and pseudouridylation have not been investigated

here, the patients affected with these mutations were found to have short telomeres and low levels of TERC. *In vitro* studies have shown that TERC is a limiting factor for telomerase activity (Wong and Collins 2006) and so these patients would be predicted to have reduced telomerase activity.

It is clear from the mutations discovered and characterised to date that the telomerase complex and telomere length is heavily implicated in DC pathology with shorter than average telomeres a common feature. Mutations have been found in five telomerase components, *DKC1* and *TERC* mutations featuring most prominently, and can be associated with a negative impact on TERC accumulation and telomerase activity. As telomerase functions to maintain telomere length it is understandable how reduced telomerase activity could lead to short telomeres. This fits well with the DC phenotype which presents most significantly in proliferating cells and tissues that are particularly susceptible to accelerated telomere shortening. Consequently, current opinion centres on DC being primarily a disorder of telomerase activity and telomere maintenance. However, it is important not to ignore the important roles dyskerin, NOP10 and NHP2 play in pseudouridylation and ribosomal processing and the potential implication for multiple defects in DC. Certainly it is noteworthy that these three H/ACA RNP components are associated with the more severe X-linked and AR forms of DC while the telomerase specific TERC and TERT are linked with the milder AD form of the disease. Perhaps the more severe phenotype is the result of a ribosomal defect also contributing to the disease? Evidence to support this can be found in mouse models of DC where *Dkc1* (*DKC1*) was mutated in an attempt to replicate this disease. Mice bearing a hypomorphic mutation that reduced *Dkc1* expression (fourfold in males, twofold in females) displayed DC like features including dyskeratosis of the skin, features of BMF and an increased propensity to developing tumours. These mice exhibited defects in pseudouridylation and rRNA processing as well as a reduction in mTR (TERC) levels and telomerase activity. However, telomere length defects were not observed until the fourth generation while the features of disease were clearly manifest from the first generation onwards (Ruggero et al. 2003). This model of DC would therefore suggest that the development of DC disease is independent of the telomerase defect. An *in vitro* mouse study specifically investigating two dyskerin mutations found in DC patients also observed defects in pseudouridylation and rRNA processing, though impacts on

telomerase activity and telomere length were only evident for one of the mutations (Mochizuki et al. 2004). Meanwhile, a study in human breast carcinomas found both TERC levels and pseudouridylation activity to be diminished in response to reduced dyskerin expression (Montanaro et al. 2006). Taken together these findings provide a convincing argument that defects in pseudouridylation and ribosomal processing could play a significant role in DC. The lack of corroborating evidence from patient studies, however, favours the disruption in telomerase activity and telomere maintenance as the principle defect in this disease (Mitchell et al. 1999b; Montanaro et al. 2002; Wong and Collins 2006).

1.7. Genetically related disorders

As already mentioned, DC has a heterogeneous clinical presentation that shares significant overlap with several other diseases. As the molecular basis of DC continues to be unravelled genetic overlap is also evident in some cases, particularly Hoyeraal Hreidarsson syndrome (HH), aplastic anaemia (AA) and myelodysplasia (MDS). HH is a severe multisystem disorder with prenatal onset that is typically characterised by intrauterine growth restriction, developmental delay, microcephaly, cerebellar hypoplasia, BMF and immunodeficiency. Investigated based on the clinical similarity to DC and x-linked inheritance in this disease, *DKC1* mutations have been identified in multiple cases of HH (Knight et al. 1999; Yaghmai et al. 2000). A homozygous *TERT* mutation has also been reported in a rare case of female HH (Marrone et al. 2007). These patients are found to have short telomeres and are considered severe variants of DC, death from BMF often occurring before the more characteristic features of DC can manifest. AA is a BMF syndrome characterised by pancytopenia and hypocellular bone marrow where the cause is idiopathic in the majority of cases. An initial presentation of AA has been associated with DC and analysis has identified *TERC* and *TERT* mutations in a subset of idiopathic AA patients that also have short telomeres (Vulliamy et al. 2002; Yamaguchi et al. 2003; Yamaguchi et al. 2005). Linking their disease to DC, these patients do not usually develop classical features of the disease and so they are considered to be ‘cryptic forms of DC’. Similarly, *TERC* mutations and short telomeres have also been identified in a subset of patients with MDS, a global BMF syndrome characterised by ineffective haematopoiesis and reduced BM progenitor cells (Yamaguchi et al. 2003).

1.8. Project aims and objectives

The mutations identified in DC patients heavily implicate the telomerase complex and telomere maintenance in the underlying molecular pathology of this disease. Initially characterised as H/ACA RNP components important for pseudouridylation and rRNA processing, the roles of dyskerin, NOP10 and NHP2 in this functionally distinct complex are less defined.

In order to better understand their implications in DC this study aims to investigate the importance of these core proteins for the integrity and function of telomerase as well as telomere maintenance. Moreover, to provide *in vitro* analysis of the potential impact dyskerin, NOP10 and NHP2 mutations may have on their function.

Finally, with growing implications for defective telomere maintenance in DC this study also aims to further investigate the prevalence of short telomeres in this disease and any potential correlations with mutation status.

Chapter 2. Materials and Methods

2.1. Materials

2.1.1. Chemicals and Reagents

All laboratory chemicals were of analytical or molecular biology grade and were obtained from BDH (supplied by VWR), Sigma-Aldrich, Fisher, or Merck. Tissue culture reagents were obtained from Lonza and Harlan. Various specialised reagents and materials used in these experiments were purchased from Invitrogen, Qiagen, Bio-Rad, Perkin Elmer or GE Healthcare, and are specifically identified in the methods section.

2.1.2. Enzymes and Kits

Various specialised kits, PCR master mixes and enzymes used in this study were supplied by Qiagen, Dako, Bio-Rad, Chemicon, Applied Biosystems, Roche, Invitrogen, Bioline, New England Biolabs (NEB), Thermo Scientific, Fermentas and Sigma-Aldrich. These are detailed individually in the methods section.

2.1.3. Primers and siRNAs

Listed in appendix B, all PCR primers and other oligonucleotides were synthesised by Sigma-Genosys and dissolved in sterile de-ionised water to give a stock concentration of 100µM, stored at -20°C. Specific TaqMan® Gene Expression Assays were purchased from Applied Biosystems as detailed in appendix B. The siRNAs used in this study are listed in appendix A and were obtained from Ambion unless otherwise stated.

2.1.4. Plasmid Expression Vectors

The pSilencer 2.1-U6 neo shRNA expression vector was purchased from Ambion and the MISSION shRNA lentiviral expression vectors from Sigma-Aldrich. The pCAG-kGP1.1R, pCAG4-RTR2, pCAGG5-VSVG and pCL10.1 MSCV Ir-GFP plasmids used for lentiviral packaging were kind gifts from Amit Nathwani (UCL). The pIRES2-EGFP plasmid was laboratory stock but originates from Clontech.

2.1.5. Bacterial Strains and Cell Lines

The DH5α ™ competent cells used for plasmid cloning were obtained from Invitrogen. The HeLa and HEK293T cell lines were laboratory stock.

2.1.6. Patient Samples

The DC patients and healthy controls included in aspects of this study were all recruited to the DCR in London, from whom samples of whole blood were obtained with written consent in accordance with the Declaration of Helsinki and with approval from the East London and City Research Ethics Committee. The DCR currently holds clinical and genetic information for 387 families worldwide and comprises 569 affected individuals (415 male, 154 female). Current criteria for entry into the DCR requires that the index case present either: (i) the diagnostic triad of mucocutaneous features (nail dystrophy, leukoplakia, and abnormal skin pigmentation), (ii) 1 or more of the mucocutaneous features (nail dystrophy, leukoplakia, and abnormal skin pigmentation), combined with a hypoplastic bone marrow and at least 2 of the other somatic features associated with DC, or (iii) at least 4 the 6 common features of HH (intrauterine growth restriction, developmental delay, microcephaly, cerebellar hypoplasia, BMF or immunodeficiency). Families of patients initially presenting AA that were found to have mutations in the *TERC* or *TINF2* genes have also be included in the DCR (Vulliamy et al. 2011b).

2.2. Methods

2.2.1. Tissue Culture

Experiments were performed using HeLa cells an immortalised human cervical cancer cell line and HEK293T cells an immortalised embryonic kidney cell line, both of which are adherent. Cells were cultured in Dulbecco's Modified eagle Medium (DMEM) containing 4.5g/l Glucose, L-Glutamine, Sodium Pyruvate (Lonza) supplemented with Foetal Calf serum to 10% (Harlen), Penicillin (100units/ml), Streptomycin (100ug/ml) and Fungizone (250ng/ml) (Lonza). Cells were cultured under humidified conditions at 5% CO_2 and 37°C. To passage the cultures all DMEM was removed and the cells washed with 1x phosphate buffered saline (PBS) (Lonza) solution. Trypsin/versin solution, enough to cover the cells, was added and incubated at 37°C for 2-3min to allow the cells to detach from the surface of the culturing flask. An equal or greater volume of DMEM was added to neutralise the trypsin/versin, the cells pelleted (1500rpm, 5min) and the supernatant removed. The cells were washed by resuspending in a suitable volume of 1x PBS and pelleted again. Depending on requirements, the cells were resuspended in a small volume of medium and counted for seeding at a specific densities or seeded less specifically i.e. 1 in 4.

2.2.2. Transfections

2.2.2.1. Transient

Cells were seeded into 6 well plates at a density of 3-5x10^5 cells per well and in 1.5ml of antibiotic free DMEM 24hr prior to transfection. Transfections were carried out using Lipofectamine 2000 or Lipofectamine RNAiMAX (Invitrogen) according to the manufactures protocol. Appropriate amounts of siRNA or plasmid were complexed with 5µl (siRNA) or 10µl (plasmid) of lipofectamine reagent in 500µl of serum and antibiotic free DMEM at room temperature for 20min before adding to the cells, giving a final culture volume of 2ml. At relevant time points, cells were harvested for analysis by detaching with trypsin/versin solution and washing as described in 2.2.1.

2.2.2.2. Stable

Cells were initially seeded into 6 well plates, cultured and transfected as described for a transient transfection. Plasmid was complexed with 10µl of lipofectamine reagent in 500µl of serum and antibiotic free DMEM at room temperature for 20min before adding to the cells giving a final culture volume of 2ml. After 24hr the cells were

transferred to a T75 flask containing 10ml of DMEM, supplemented after 48hr with neomycin (G418) at 800µg/ml and maintained for 2-3 weeks to select for stably transfected cells conferring the antibiotic resistance. After this time neomycin was used at 400µg/ml to maintain the selective pressure.

2.2.3. Lentiviral Packaging

The human embryonic kidney cell line HEK293T was used to package lentiviral particles for downstream transduction of target cells. Cells were seeded into T175 flasks and cultured in antibiotic free DMEM to between 50-75% confluence. Plasmid constructs expressing the necessary components for lentiviral packaging were then transfected into the cells using Lipofectamine LTX (Invitrogen) according to the manufactures protocol. 12µg gag/pol enzymatic component plasmid: pCAG-kGP1.1R, 4µg Rev/Tat response element packaging plasmid: pCAG4-RTR2, 4µg vesicular stomatitis virus G protein envelope plasmid: pCAGG5-VSVG and 20µg lentiviral genome vector: pCL10.1 MSCV Ir-GFP (All gifts from Amit Nathwani, UCL (Kyriakou et al)) were mixed and then complexed with 94µl of Lipofectamine LTX reagent in 7.5ml of antibiotic and serum free DMEM at room temperature for 30min. The transfection complexes were then applied to the cells, replacing the medium in the flask with 22.5ml of fresh antibiotic free DMEM to give a final culture volume of 30ml. After culturing overnight the cells were rinsed in the flask with 1x PBS twice before adding 16ml of fresh DMEM. Cells were then cultured for a further 24-36hr while they produced packaged lentiviral particles. The lentivirus containing culture medium was then collected and passed through a 0.45µm filter to remove any cells.

2.2.4. Lentiviral Concentration

Lentiviral particles were concentrated by complexing with chondroitin sulphate C (CSC) and polybrene (PB). Filtered lentivirus containing medium was divided into 1.5ml aliquots then PB (Sigma) and CSC (Sigma) added, 4µl (20mg/ml) each. Samples were incubated at 37°C for 20min to allow lentiviral complexes to form before pelleting them (13000rpm, 5min). Each lentivirus containing pellet was resuspended in 100µl of 1x PBS before pooling together and separating into 50µl aliquots for use in downstream applications or storage at -70°C.

2.2.5. Lentiviral Transduction

Cells were seeded into a 6 well plate at a density of 3-5x10^5 cells per well and in 2ml of DMEM 24hr prior to transduction. Viral particles were applied to the cells in 50µl aliquots and cultured for 6hr before transferring the cells to a T75 flask containing 10ml of DMEM. Culture medium was supplemented with puromycin to 1µg/ml after 48hr and maintained for a further 2-3 weeks to select for transduced cells conferring the puromycin resistance.

2.2.6. RNA Extraction and DNase Treatment

Total RNA was extracted from cells using the RNeasy mini kit (Qiagen) according to manufacturer protocols. Cell pellets were resuspended in 350µl of the RLT cell lysis buffer (Qiagen), homogenized and stored at -70°C to be completed in larger batches at a later date. To complete the extraction cell lysates were precipitated with an equal volume of 70% ethanol and applied to RNA binding spin columns (13,000rpm, 30sec). Columns were subsequently washed through once with 700µl of buffer RW1 (Qiagen) and then twice with 500µl of buffer RPE (70% ethanol) (Qiagen) before 'dry' spinning the column to ensure all residual buffer was removed (all spins 13,000rpm, 1min). RNA was eluted from the column into ~35µl of RNase/DNase free water. DNA contamination was removed by adding 1 unit of DNase1 enzyme (Invitrogen), 4µl of 10x DNase buffer and incubating at 37°C for 15min. The enzyme was then deactivated by heating to 65°C for 10min. All RNA samples were stored at -70°C.

2.2.7. First Strand cDNA Synthesis

First strand cDNA synthesis was performed using M-MLV RT enzyme (Invitrogen) according to manufacturer protocols. Total RNA was used to prepare cDNA in 40µl reactions consisting of 20µl RNA (DNase treated), 1x reaction buffer (50mM Tris-HCl pH 8.3, 75mM KCl, 3mM $MgCl_2$), 10mM DTT, 60units RNasin (Bioline), 200ng random hexamers (Qiagen), 0.5mM dNTP mix (Invitrogen), 300units M-MLV RT enzyme. The reaction mixture was incubated at 37°C for a minimum of 1hr before the reaction was stopped by heat inactivating the RT enzyme at 70°C for 15min. All cDNA samples were stored at -20°C.

2.2.8. Standard PCR

PCR was performed using *Taq* DNA polymerase (Thermo Scientific) in a programmable thermocycler (Eppendorf Mastercycler). 25µl reactions consisted of 1µl cDNA or DNA template, 1x PCR buffer (Thermo Scientific; 75mM Tris-HCl pH 8.8, 20mM $(NH_4)_2SO_4$, 0.01% Tween® 20), 1.5mM $MgCl_2$, 0.2mM dNTP mix (Invitrogen), 0.2mM of each primer (forward and reverse) and 1 unit of enzyme. Reactions were run using a standard PCR cycle program: 95°C for 2min; then 35 cycles of 95°C for 30sec, 58°C for 45sec, 72°C for 1min; followed by 72°C for 5min.

2.2.9. Quantitative Real Time PCR

Quantitative real time PCR (qPCR) uses sequence detection methods to quantify amplification products after each cycle of the PCR. One approach uses dual-labelled oligonucleotide probes specific to a short sequence within that to be amplified. During amplification the exonuclease activity of the *Taq* polymerase releases a fluorophore from the probes 5' end which allows it to fluoresce by liberating it from the effects of a quencher at the 3' end. Alternatively, DNA binding dyes can be used that only fluoresce when intercalated with double stranded DNA i.e. amplified PCR products. In both cases the fluorescence generated is directly proportional to the amount of amplified products and is measured after each cycle allowing product accumulation to be quantified. Quantification is achieved by reference to either a calibrator sample (relative quantification) or to a standard curve generated using data from samples in which the copy number is already known (absolute quantification). To control for variations in sample loading values are normalised to that of an endogenous control that is also quantified. In this study both methods of detection and quantification have been employed using the endogenous *ABL* gene to normalise.

2.2.9.1. Standard Curve Preparation

Standard curve samples were generated from PCR products for the target gene to be quantified, initially amplified in a standard PCR reaction. A post amplification copy number was determined using product concentration, length, the molecular weight of DNA (660g/bp) and Avogadro's constant (6.022E+23) in the following calculation.

$$\frac{\text{Concentration (ng/ul)} \times 6.022E+23}{\text{Length (bp)} \times 1x10^9 \times 660}$$

The PCR products were then serially diluted to generate an 8 point standard curve comprising 10^9, 10^7, 10^6, 10^5, 10^4, 10^3, 10^2 and 0 copies/2µl.

2.2.9.2. Probe Based Detection – TaqMan® Probes

Reactions were carried out using the TaqMan® probe PCR system (Applied Biosystems) according to manufacturer protocols. 20µl reactions consisted of 2µl cDNA, 1x TaqMan® Universal Master Mix (Applied Biosystems) and primers and probes that varied depending on the gene assayed (details in appendix B). Reactions were set up in triplicate and run on an ABI 7500 real time PCR machine (Applied Biosystems) using the following cycle program: 50°C for 2min, 95°C for 10min; then 50 cycles of 95°C for 15sec, 60°C for 1min* (* denotes fluorescence signal detection step).

2.2.9.3. DNA Dye Based Detection – SYBR® Green

Reactions were carried out using SYBR® Green PCR Master Mix, (Applied Biosystems) according the manufacturers protocol. 20µl reactions consisted of 1µl cDNA, 1x SYBR Green PCR Master Mix (Applied Biosystems), and 0.2mM of each primer (details in appendix B). Reactions were set up in duplicate and run on an ABI 7500 real time PCR machine (Applied Biosystems) using the following cycle program: 50°C for 2min, 95°C for 10min; then 50 cycles of 95°C for 15sec, 60°C for 1min* (* denotes fluorescence signal detection step).

2.2.9.4. qPCR Analysis

Initial analysis of the raw fluorescence data generated was performed in an automated process using the 7500 System Sequence Detection Software v1.2.3 (Applied Biosystems). Firstly, fluorescence signals were normalised to that of the passive reference dye ROX™, included within the PCR master mix, to give a reporter normalised (Rn) value for each cycle of the reaction. Next, a baseline signal was determined to account for background fluorescence and the Rn values baseline corrected to give ΔRn values. Finally, amplification curves were plotted using log (ΔRn) against cycle number and a threshold ΔRn value set to intersect the exponential (or geometric) phase of the curves. The point of intersection determines the threshold cycle (C_T) number which is used as the basic measure of relative target concentration

within the PCR. The C_T values for experimental and control samples were then used in subsequent calculations to achieve either absolute or relative target quantification.

2.2.9.4.1. Absolute Quantification

Absolute quantities were determined by comparison to a standard curve, generated by plotting log copy number against C_T for the standard curve samples. The equation of the line was then used to extrapolate copy number information from the C_T values of unknown samples. Copy numbers were calculated for each PCR replicate before averaging to obtain a final copy number for the sample. Target copy numbers were then normalised to those of the *ABL* endogenous control also measured in each sample as a reference. This was achieved by dividing the target and *ABL* copy numbers to give a target/*ABL* ratio which is used as the final measure of target quantity for comparison.

2.2.9.4.2. Relative Quantification

Relative quantities were determined using the comparative C_T method described by Applied Biosystems. The amount of target was normalised to the *ABL* reference and calculated relative to a designated calibrator sample using only C_T values in the mathematical formula $2^{-\Delta\Delta C}{}_T$ as follows.

1. *ABL* (R) C_T subtracted from target (X) C_T to give ΔC_T.

$$\mathbf{\Delta C_T = C_{T,X} - C_{T,R}}$$

2. Calibrator (C) ΔC_T subtracted from sample (S) ΔC_T to give $\Delta\Delta C_T$.

$$\mathbf{\Delta\Delta C_T = \Delta C_{T,S} - \Delta C_{T,C}}$$

3. Relative quantity = $\mathbf{2^{-\Delta\Delta C}{}_T}$

2.2.10. Agarose Gel Electrophoresis

PCR products and other nucleic acids samples were separated by electrophoresis on agarose gels made at 1.5% in 100ml of 0.5x Tris-Borate-EDTA buffer (TBE: 90mM Tris, 90mM Boric acid, 2mM EDTA), with 10µg of ethidium bromide. 6x loading buffer (3x TBE, 30% glycerol, 0.3% bromophenol blue) was added to the samples to give 1x final concentration and 10µl loaded for electrophoresis. Gels were run in 0.5x TBE buffer and visualised under UV illumination.

2.2.11. Gel Extraction and Purification

Gel extraction and purifications were carried out using the QIAquick Gel Extraction kit (Qiagen). DNA bands were cut from agarose gel under UV illumination. A volume of Buffer GE (Qiagen) was added corresponding to the 3x the weight of the excised gel fragment (100µg = 300µl) and incubated at 50°C for 10min to dissolve the gel. A volume of 100% isopropanol corresponding to 1x the weight was added before applying to a DNA binding spin column (all spins 13,000rpm, 1min). Columns were subsequently washed through with Buffer PE (Qiagen) and the DNA eluted into 30-50µl of buffer EB (Qiagen). Purified DNA bands stored at -20°C.

2.2.12. Oligonucleotide Annealing

Long, complementary oligonucleotides encoding shRNA sequences were annealed to create double stranded DNA molecules for cloning into the p*Silencer*™ 2.1-U6 neo expression plasmid (Ambion). Oligonucleotides were diluted to 1µg/µl in TE buffer (10mM Tris, 1mM EDTA) prior to annealing. In 50µl reactions, 2µg of each oligonucleotide were mixed in 1x DNA annealing solution (Ambion) and heated to 95°C for 3-4min before leaving to cool slowly for 1hr. The annealed oligonucleotides are then ready for downstream applications or storage at -20°C.

2.2.13. Ligation

Ligations were carried out using T4 DNA ligase (Fermentas) in 10µl reactions consisting of 1µl annealed oligonucleotides, 1µl linear p*Silencer*™ 2.1-U6 neo (Ambion), 1µl ligase mixed in 1x ligation buffer (Fermentas). Reactions were incubated at 4°C overnight and stopped by heating to 65°C for 10min.

2.2.14. Transformation of Competent *E.Coli*.

Transformations were performed using DH5α ™ competent cells (Invitrogen), 25µl of the cells were mixed with 3µl of ligation reaction and incubated on ice for 30min. Reactions were then heat shocked at 42°C for 20sec and returned to the ice for a further 2min before adding 500µl of Luria-Bertani broth (LB) (Merck) and incubating at 37°C for 1hr. 70µl were then plated on LB agar (Merck) plates containing ampicillin (Sigma) at 50µg/ml and incubated at 37°C overnight.

2.2.15. Small Scale Plasmid DNA Preparation

Bacteria from a single colony were cultured overnight in 2ml LB, supplemented with ampicillin at 50µg/ml, at 37°C with shaking. Cells were then pelleted (13,000rpm, 5min) and plasmid extracted using the QIAprep Spin Miniprep kit (Qiagen) according to manufacturer's protocols. Cells were resuspended in 250µl of Buffer P1 (Qiagen) before lysing with 250µl of Buffer P2 (Qiagen), after 5min the lysis reaction was stopped with 350µl of Buffer N3 (Qiagen). Protein precipitate was pelleted (13,000rpm, 10min) and the supernatant applied to the DNA binding spin columns (all spins 13,000rpm, 1min). Columns were subsequently washed through with 750µl of Buffer PE (Qiagen), spun empty to remove any residual buffer and the plasmid DNA eluted from the column with 50µl of buffer EB (Qiagen). Plasmids were stored at -20°C.

2.2.16. Large Scale Plasmid DNA Preparation

Bacteria from a single colony were cultured for 6-8hr in 1ml LB supplemented with ampicillin to 50µg/ml, at 37°C with shaking. The starter culture was then added to 500ml LB supplemented with ampicilin to 50µg/ml and cultured overnight at 37°C with shaking. Bacteria were then pelleted (6000x *g*, 15min, 4°C) and plasmid DNA extracted using the QIAGEN Plasmid Maxi kit (Qiagen) according to manufacturer protocols. Bacterial cells were resuspended in 10ml of Buffer P1 (Qiagen) before lysing with 10ml of Buffer P2 (Qiagen) at room temp for 5min. The lysis reaction was stopped with 10ml of chilled Buffer P3 (Qiagen), incubating on ice for 15min to precipitate the protein. The precipitate was then pelleted (20,000x *g*, 30min, 4°C) and the plasmid DNA containing supernatant centrifuged a second time (20,000x *g*, 15min, 4°C) to ensure all protein precipitate was removed. The supernatant was then applied to DNA binding filter columns (QIAGEN-tip 500), pre-equilibrated with 10ml of Buffer QBT (Qiagen), and allowed to pass through by gravity flow. Columns were subsequently washed through twice with 30ml Buffer QC (Qiagen) and the plasmid DNA eluted from the column with 15ml of buffer QF (Qiagen). Plasmid DNA was then precipitated by mixing with 10.5ml of isopropanol, pelleted (15,000x *g*, 30min, 4°C) and washed with 5ml of 70% ethanol before pelleting again (15,000x *g*, 10min). The plasmid DNA pellet was left to air dry for 5-10min before dissolving in a suitable volume (typically 1ml) of TE buffer (Qiagen). Plasmids were stored at -20°C.

2.2.17. Genomic DNA Extraction

Genomic DNA was extracted from cells using the Gentra® Puregene® DNA extraction kit (Qiagen) according to manufacturer protocols. Cell pellets were resuspended in a suitable volume of Cell Lysis Solution (Qiagen) to completely lyse the cells. Protein was then removed from the lysate by adding a volume of Precipitation Solution (Qiagen) (equivalent to 1/3 of the volume used to lyse the cells) and incubated on ice for 5min. Precipitate was pelleted (13000rpm, 1min) and the DNA containing supernatant carefully removed to a clean tube. DNA was then precipitated by adding a volume of 100% isopropanol (equivalent to that used to lyse the cells) and pelleted (13000rpm, 1min). The DNA pellet was washed in 300µl of 70% ethanol and pelleted again (13000rpm, 1min). The ethanol wash was carefully removed and the DNA pellet left to air dry for 5-10min before resuspending in a suitable volume of DNA Hydration Solution (Qiagen). Pellets were either left to slowly dissolve overnight or more rapidly by heating to 60°C for 1hr.

2.2.18. DNA/RNA Quantification

DNA and RNA samples were quantified using the NanoDrop 1000 spectrophotometer (Thermo Scientific). 1µl of neat sample was applied directly to the machine which directly measured and calculated concentration as ng/µl.

2.2.19. Telomere Length Measurement by Southern Blot

Employing a subtelomeric probe specific to chromosome 7 (pTelBam8) (Brown et al. 1990), southern blotting was used to specifically identify and measure telomere lengths within a DNA sample. Genomic DNA was digested completely with a restriction endonuclease to create smaller restriction fragments. These were subsequently size separated using agarose gel electrophoresis before transfer and immobilisation onto a nylon membrane. The DNA was fixed on the membrane before hybridising with a radio-labelled DNA probe to identify the telomeric DNA.

5-10µg of genomic DNA was digested with the *BamHI* restriction enzyme in 70µl reactions consisting of 4µl *BamHI* (NEB), 7µl 10x RE buffer (NEB) and incubated at 37°C overnight (16hr +/-). Loading buffer (see 2.2.10) was added to a 1x final concentration and the digested products separated by electrophoresis on 0.8% agarose

gels made with 450ml 1x Tris-Acetate-EDTA buffer (TAE: 40mM Tris, 20mM Acetic acid, 1mM EDTA) containing 150µg of ethidium bromide. Gels were run at ~60v in 1x TAE buffer for 20-24hr to allow sufficient separation of the restriction fragments. Under UV illumination to enable visualising the DNA, excess areas were trimmed and discarded prior to soaking the gel in an alkaline denaturing solution (1.5M NaCl, 0.5M NaOH) for 40min to denature the DNA. The gel was then assembled into a 'blotting' stack to facilitate transfer of the DNA from the gel to the nylon membrane. This was done by placing the gel, face down, onto a larger sized piece of 20mm whatman paper, the 'wick', before carefully layering on top a corresponding sized piece of nylon Hybond N+ membrane (GE healthcare) and 2 size matched pieces of 20mm whatman paper. Extra care was taken not to create or to remove any air bubbles that formed between the different layers. Paper towelling was then stacked on top to a height of 20-30cm. The stack was assembled above a tray containing 20x saline sodium citrate solution (SSC: 3M NaCl, 0.3M trisodium citrate, pH 7) into which the ends of the wick extended. This was left overnight (16hr +/-) to allow SSC to be absorbed through the stack and to transfer of ssDNA from the gel to the membrane. The membrane was then rinsed in 2 changes of 3x SSC before incubating at 80°C for 2-3hr in order to cross link the DNA with the membrane. After briefly rehydrating in 3x SSC the membrane was incubated in 40ml pre-hybridisation solution (6x SSC; 2.5x Denhardt's solution: 0.05% Ficoll 400, 0.05% PVP-40, 0.05% BSA; 10mg carrier DNA and 0.75% SDS) for 4hr at 67°C in preparation for hybridisation with the probe.

^{32}P labelled DNA probes were prepared using the Megaprime kit (GE healthcare). Primer and probe DNA were pre-annealed in 35µl reactions consisting 5µl primer solution and 20-30ng probe DNA heated to 95°C for 5min before slow cooling on the bench. This was then supplemented to also consist of 1x labelling buffer (GE healthcare), 2-5µl [α^{32}P] dCTP (6000Ci/mmol, 20mCi/ml) (Perkin Elmer) and 2µl klenow fragment DNA polymerase, making a final volume of 50µl, before incubating at 37°C for 30min. The labelled probe was then purified by adding 50µl TE + 0.1% SDS and applying to a sephadex G50 spin column (1700x *g*, 4min) before mixing with 400µl of carrier DNA (10µg/µl). Labelled probe was heat denatured at 95°C for 5min before incubating with the membrane in 10ml hybridisation solution (6x SSC,

2.5x Denhardt's solution, 10mg carrier DNA and 0.75% SDS) at 67°C overnight (16hr +/-). The membrane was then subjected to a series of washes to remove any unbound probe, all at 67°C as follows: 10ml 2x SSC for 10min; 50ml 2x SSC for 30min; twice in 10ml 0.2x SSC for 10min. The membrane was wrapped in plastic and secured in a developing cassette for exposure overnight (16-24hr +/-) or longer.

Note - All radioactive aspects of this protocol were carried out by Dr T Vulliamy.

2.2.20. Telomere Length Measurements by Flow FISH

Adapting the protocol according to Baerlocher et al the Telomere PNA kit/FITC (Dako) was used to measure relative telomere length (RTL) within nucleated white blood cells (Baerlocher et al. 2006). The assay involves fluorescence in situ hybridisation (FISH) performed on a cell suspension consisting of white blood cells mixed with calf thymocytes (as a control cell) both in the presence and absence of a telomere specific PNA/FITC probe which can be detected by flow cytometry. Measured a single cell at a time, the fluorescence generated is proportional to the total amount of telomeric DNA within the cell. Using samples treated without the probe to control for background cell fluorescence, an average RTL for the white cell population can be determined using a calibrator sample of pre-determined telomere length.

2.2.20.1. White Blood Cell Preparation

White blood cells were separated from samples of whole blood using red cell lysis buffer (RCLB; 155mM NH_4Cl, 10mM $KHCO_3$, 0.1mM EDTA, pH 7.4). A volume of whole blood was mixed with a minimum of 3 volumes of ice cold RCLB and incubated at room temperature for a minimum of 10min to lyse the red blood cells. White blood cells were then pelleted (2000rpm, 5min) and the supernatant carefully discarded. The cell pellet was resuspended in 1ml of ice cold RCLB and transferred to a microfuge tube before centrifuging briefly (12000rpm, 15sec) to pellet the cells again. The supernatant was carefully removed and the white cell pellet resuspended in a suitable volume of PBS, typically 300µl, and mixed with an equal volume of cell freezing medium (FCS, 20% DMSO) before storage at -70°C.

2.2.20.2. Probe Hybridisation

Approximately $2x10^6$ calf thymocytes were mixed with the white blood cell sample and then pelleted (12000rpm, 15sec). The supernatant was carefully discarded and the cells resuspended in 1ml of PBS before dividing into duplicate 500µl aliquots and pelleting again (12000rpm, 15sec). Carefully discarding the supernatant, one of the duplicates was resuspended in 150µl of Hybridization Solution (Dako) and the other in 150µl of Telomere PNA Probe/FITC in Hybridization Solution (Dako) before incubating at 82°C for 10min to denature the DNA. Samples were mixed briefly on a vortex and left in a dark place at room temperature to cool overnight and allow the probe to hybridise. Samples were washed by mixing with 500µl of Wash Solution (Dako) and incubating at 40°C for 10min before pelleting the cells (1500x *g*, 5min, 16°C). The supernatant was carefully removed, the cells resuspended in 500µl of Wash Solution and incubated at 40°C for 10min for a second time before pelleting again (900x *g*, 5min, 16°C). The supernatant was carefully removed and the cells resuspended in 300µl of DNA counterstaining solution (PBS, 0.1% BSA, 1000 U/ml Rnase T1, 12.4ng/ml LDS751) before incubating at 4°C for a minimum of 30min. Samples were then ready for analysis by flow cytometry and could be stored at 4°C for up to 7 days.

2.2.20.3. Flow Cytometry – Data Acquisition and Analysis

Samples were analysed on a BD FACSCanto II (BD Biosceinces) flow cytometer using BD FACSDiva™ software v6.1.3 (BD Biosceinces). LDS751 DNA stain was detected using the PerCP-Cy5-5 (red) channel and the PNA probe using the FITC (green) channel, both set to linear. Adapted from Baerlocher et al, cell populations were selectively gated and data acquired for a combined total of 30-50,000 events as outlined in figure 2.1 A-C. All data was collected from the SSC, FSC, PerCP-Cy5-5 and FITC channels for later analysis. FlowJo software v.8.7.1 (Tree Star) was used to perform post acquisition analysis in which the cell populations were gated again and median FITC values obtained for each of the cell types as outlined in figure 2.1 A-E.

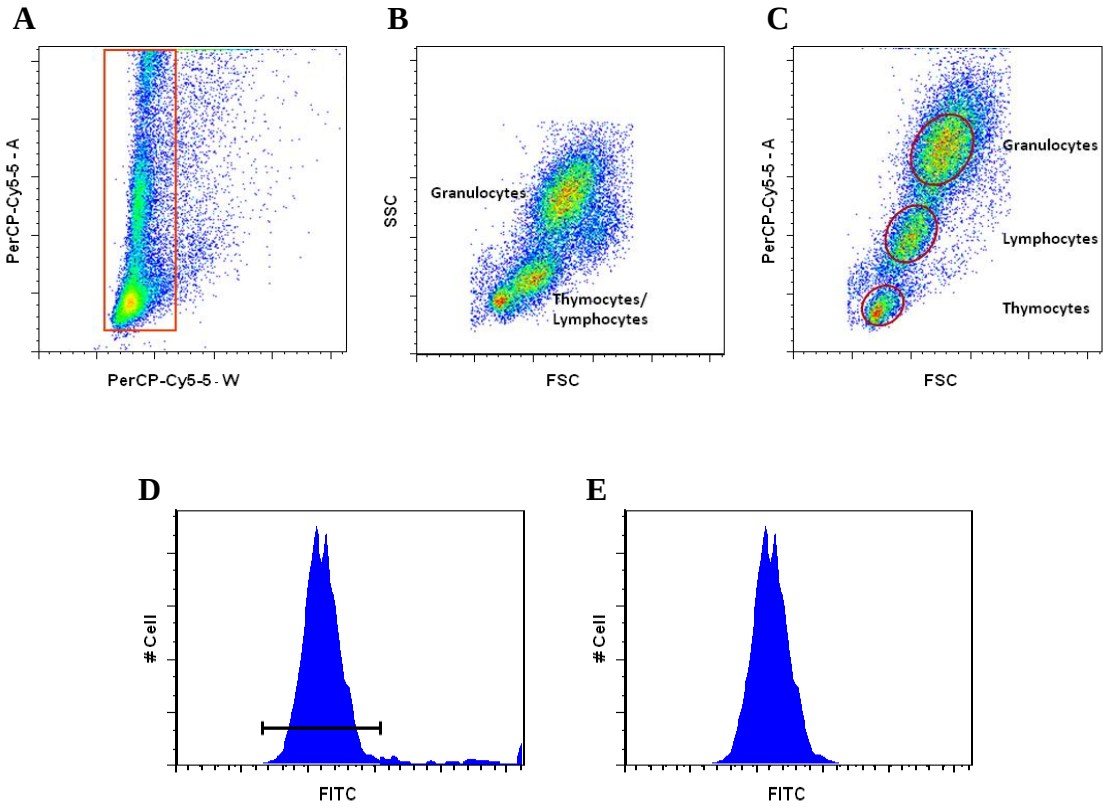


Figure 2.1. RTL by flow FISH. (A) Single cells were gated using PerCP-Cy5-5-A against PerCP-Cy5-5-W. (B) Sample quality was verified using SSC against FSC, checking for the mixed thymocyte/lymphocyte and granulocyte populations. (C) Separation of the thymocytes and lymphocytes populations was achieved using PerCP-Cy5-5-A against FSC, allowing the 3 cell types to be gated individually. (D) For each, FITC singles were visualised as a histogram and any 'low level noise' removed by gating the predominant population. (E) Median FITC values were obtained for the final gated populations.

2.2.20.4. Calculating Relative Telomere Length (RTL)

For each sample delta FITC median values were calculated by subtracting the median obtained from treatment without probe (background) from those obtained with probe, for each of the cell populations: thymocytes, lymphocytes and granulocytes. The delta FITC median values were then used in the subsequent calculation to determine the telomere length relative to the calibration sample.

$$\left(\frac{A_X}{B_X} \times \frac{B_c}{A_c} \right) \times ?kb$$

A = Cell type in question (Lymph or Gran)
B = Thymocytes
C = Calibration sample
X = Sample in question
? = Telomere length of calibration sample

(Taken from Baerlocher et al 2006)

2.2.21. Telomere Length Measurement by Monochrome Multiplex qPCR

Monochrome multiplex qPCR (MMqPCR) was used to measure relative telomere lengths within a DNA sample. Devised by Cawthon, this technique is a manipulation of qPCR to determine telomere length from a DNA sample (Cawthon 2009). From genomic DNA a telomere specific amplicon (T) is amplified by qPCR using the SYBR Green DNA dye for detection and quantification. The amplicon is of a fixed length (79bp) but is able to prime at sites along the full length of the telomere. As a result longer telomeres are able to produce more amplicons meaning that the amount of T generated is proportional to telomere length. Therefore, quantification of T can be used as a measure of the average telomere length within the DNA sample. To control for variations in DNA loading, T is normalised to a single copy gene amplicon (S) also quantified, in this case the human beta globin gene. Here, T and S values are divided to give a T/S ratio that is used as the final measure of average telomere length for the DNA sample. Quantification is achieved by reference to relative standard curves for T and S generated by titration of total DNA input of a chosen ‘standard DNA’. Consequently, T/S ratios for the unknown samples are relative to that of the standard DNA, which is always 1. Therefore, samples with a T/S > 1.0 have an average telomere length that is relatively longer than the standard DNA while samples with a T/S < 1.0 have an average telomere length that is relatively shorter.

2.2.21.1. MMqPCR Reactions

DNA samples were carefully diluted to 2ng/µl and the standard DNA to 15ng/µl, serially diluted to give a five point standard curve of 15, 5, 1.666, 0.555, 0.185ng/µl. 15µl reactions consisted of 1x Roche SYBR Green Master I reaction mix (Roche Applied Science), T primers at 1µM each (telc and telg, appendix B), S primers at 200nM each (hbgu and hbgd, appendix B), and 5µl of DNA. This gave a total DNA input of 10ng for experimental samples and 75, 25, 8.333, 2.775, 0.925ng for the five point standard curve. Reactions were performed in triplicate and run on a Roche LightCycler 480 real time thermal cycler (Roche Applied Science) using cycle parameters of: 95°C for 15min; 2 cycles of 94°C for 15sec, 49°C for 15sec; and then 32 cycles of 94°C for 15sec, 62°C for 10sec, 74°C for 15sec*, 84°C for 10sec, 88°C for 15sec* (* denotes fluorescence signal detection steps).

2.2.21.2. MMqPCR Analysis

The multiplex nature of MMqPCR requires fluorescence data to be acquired twice during each cycle. This generates two sets of fluorescence data, one for T and one for S, that have to be separated prior to analysis in order to generate independent amplification curves for the two. Raw data was exported from the LC480 lightcyler as a 'txt' file and imported into excel where the two data sets, identifiable by cycle details, could be separated manually. Individual data files for T and S were then passed through the file conversion software Conversion v1.5 (Dr J.M.Ruijter) from which they could be imported into the software LinRegPCR v12.2 (Dr J.M.Ruijter) for analysis. In an automated process, the raw fluorescence data was normalised, baseline corrected and plotted to obtain C_T values as described in 2.2.9.4.

Standard curves for T and S were plotted as log DNA input against C_T and the equation of the line used to extrapolate DNA input information (as a measure of relative abundance) from the C_T values of unknown samples. DNA input was calculated for each PCR replicate before averaging to obtain a final DNA input value. T and S values from the same sample were then divided to give a T/S ratio used as the final measure of relative telomere length.

2.2.22. Protein Extraction

Whole cell protein extracts were prepared using Radio-ImmunoPrecipitation Assay (RIPA) cell lysis buffer (50mM Tris-HCl pH 8, 150mM NaCl, 1% NP40, 0.1% SDS, 0.5% sodium deoxycholate) containing 1mM PMSF (phenylmethylsulphonyl fluoride) and cocktail protease inhibitors (Sigma). Cell pellets were lysed by mixing with a suitable volume of freshly prepared ice cold RIPA lysis buffer and incubating on ice for 30mins. Cellular debris was pelleted (12,000rpm, 25min, 4°C) and the protein containing supernatant transferred to a fresh tube. Protein lysate were stored at -70°C.

2.2.23. Protein Assay

Protein concentrations were measured using the *DC*™ protein assay (Bio-Rad), a detergent compatible colorimetric assay for soluble protein. Based on the Lowry assay, the protein sample is treated with alkaline copper tartrate and Folin reagent sequentially. The alkaline copper treated proteins remove oxygen atoms from the Folin reagent in a reduction reaction that results in a quantifiable blue colour change proportional to the protein concentration. Comparison to a standard curve enables the amount of protein to be calculated. Reactions were performed according to the manufacturer's microplate assay protocol. Protein standard II (bovine serum albumin) (Boi-Rad) was used to create a standard curve by serially diluting from 1.6 – 0.1mg/ml. 5µl of protein lysate or standard were first mixed with 25µl of reagent A (Bio-Rad), then 200µl of reagent B (Bio-Rad) and incubated at room temperature for 15mins. Colour change was then measured by absorbance at 750nm and protein concentration calculated against the standard curve.

2.2.24. Telomerase Repeat Amplification Protocol (TRAP) Assay

Telomerase activity was measured using the TRAPeze® RT Telomerase Detection Kit (Chemicon). This is a fluorometric PCR based assay for *in vitro* analysis of telomerase activity in the cells of interest. The assay initially involves preparing a cell extract which is subsequently used in two separate enzymatic reactions performed in one tube. In the first, telomerase activity extends a synthetic telomere (TS) mimicking telomere extension within the cell. In the second, the extended TS is amplified by qPCR using the fluorescently labeled Amplifluor® reverse primer complementary to GGTTAG of the telomere repeat sequence. Fluorescence emission occurs only when

the Amplifluor® reverse primer is incorporated into a PCR product, the measure of which is directly proportional to the amount of TRAP products generated i.e. to the level of telomerase activity in the first reaction. Quantification is achieved by reference to a standard curve which is generated using the TSR8 control template (TS primer extended with 8 telomeric repeats AG$(GGTTAG)_7$).

2.2.24.1. TRAP Assay Reactions

Reactions were setup and performed according to the manufacturer protocol. Cell extracts were prepared by lysing with ice cold CHAPS lysis buffer (Chemicon), typically 150-200µl, and incubating on ice for 30mins. Cellular debris was pelleted (12,000rpm, 25min, 4°C) and the protein containing supernatant transferred to a fresh tube. Protein concentration was determined and a 500ng/µl dilution made for use in the actual TRAP assay. Assay controls included heat inactivated extracts, where an aliquot of cell extract was heated to 85°C for 10min to inactivate telomerase, and CHAPS buffer alone. A four point standard curve was prepared by diluting the TSR8 control template to 20, 2, 0.2, and 0.02amoles/µl. 25µl reactions consisted 2µl of either cell extract (1µg), TSR8 dilutions or control sample, 1x TRAPeze® RT reaction mix (Chemicon) and 2 units of titanium *Taq* (Clontech). Reactions were run on an ABI 7500 real time PCR machine (Applied Biosystems), first incubating at 30°C for 30min before entering into the following PCR program: 95°C for 2min then 45 cycles of 94°C for 15sec, 59°C for 1min, 45°C 30sec* (* denotes fluorescence signal detection step).

2.2.24.2. TRAP Assay Analysis

As a qPCR based assay data analysis was performed as described for absolute quantification by qPCR in sections 2.2.9.4 and 2.2.9.4.1 with one exception. The TRAP assay does not include quantification of a reference against which to normalise for sample loading variation. Therefore copy number is the final measure of quantity for subsequent comparisons.

2.2.25. Western Blotting

Specific proteins were identified in cell lysates by western blotting under reducing and denaturing conditions. Proteins were size separated by polyacrylamide gel electrophoresis (PAGE) before transfer and immobilisation onto a nitrocellulose

membrane. The membrane was then blocked with a milk solution and probed with a primary antibody specific to the protein under investigation. Detection was achieved using horseradish peroxidise (HRP)-conjugated secondary antibodies specific to the primary antibody and enhanced chemoluminescence (ECL) reagents, a mixture of luminol and hydrogen peroxide. In the presence of hydrogen peroxide HRP catalyses the oxidation of the luminol exciting it to emit light which can be detected using X-ray film.

PAGE gels were prepared at the desired concentration (typically 10%) composed of a lower running gel (0.375M Tris-HCl pH8.8, 0.2% SDS, 10% polyacrylamide, 0.033% ammonium persulfate, 0.066% TEMED; typically 6-7ml final volume) and an upper stacking gel (0.25M Tris-HCl pH6.6, 0.4% SDS, 2.6% polyacrylamide, 0.033% ammonium persulfate, 0.066% TEMED; typically 3-4ml final volume) in which the sample wells were set. Protein samples (typically 10µg) were prepared by adding loading buffer to a 1x concentration (62.5mM Tris-HCl pH6.8, 10% glycerol, 2% SDS, 5% DTT, 0.002% bromophenol blue) and heat denaturing at 100°C for 5-10min. Samples were loaded alongside a size standard and gels run for 90min at 100v in 1x running buffer (25mM Tris, 0.2M glycine, 0.1% SDS). Post electrophoresis the upper stacking gel was trimmed from the lower running gel and discarded. To transfer proteins from the gel, a corresponding sized piece of nitrocellulose membrane (Bio-Rad) was carefully laid upon the gel which was then sandwiched between pieces of 20mm whatman paper of the same size, 3 pieces on either side. Extra care was taken not to create or to remove any air bubbles that formed between the different layers. Proteins were transferred from the gel to the membrane for 1hr at 100v in 1x transfer buffer (25mM Tris, 0.2M glycine) after which the membrane was incubated in 10ml blocking solution (1x Tris buffered saline (TBS): 50mM Tris-HCl pH7.4, 150mM NaCl; 5% powdered milk) for at least 1hr to reduce non specific antibody binding. An appropriate dilution of the primary antibody was prepared in blocking solution and added to the membrane for 2-3hr at room temperature or overnight at 4°C. The membrane was washed 3 times in 10ml 1x TBS for 5min before incubating with an appropriate dilution of secondary antibody, also in blocking solution, for 2hr at room temperature. The membrane was washed 3 times in 10ml 1x TBS, 0.5% Tween for 5min before developing in 5ml of ECL solution (Thermo Fisher) for 5min. The

membrane was wrapped in plastic and secured in a developing cassette for X-ray film exposure.

2.2.26. Sequencing

Aliquots of samples and primers were sent for external sequencing at the QMUL core genomics centre. Sequencing reactions were performed by dideoxy nucleotide termination utilising Big Dye™ chemistry (Applied Biosystems) with products separated by capillary electrophoresis. The sequence data returned was viewed and aligned using VectorNTI v10 software (Invitrogen).

2.2.27. Statistical Analysis

The Mann-Whitney U test was used to determine the statistical significance of differences between healthy controls and DC patient groups in some aspects of this study. Performed using GraphPad Prism 5 software (GraphPad Software Inc), this tests the null hypothesis that observations from two populations are similar against the alternative hypothesis that those from one are larger than the other by comparing the sum of their ranks. Because of this the Mann-Whitney U test does not require the populations to be normally distributed and provides a non-parametric alternative to the unpaired t-test that is more robust in the presence of outliers. The calculation for this test is outlined below.

$$U = N_1 N_2 + \frac{N_1(N_1 + 1)}{2} - R_1$$

N is the sample size, R is the sum of ranks

Chapter 3. Transient Knockdown of Dyskerin, NOP10, NHP2 and GAR1

3.1. Introduction

As discussed in chapter 1, impaired telomere maintenance features prominently in DC with different patients found to harbour mutations affecting several telomerase components including the proteins dyskerin, NOP10 and NHP2. Together with GAR1 these form a protein core that is common to telomerase and all other H/ACA RNPs (Mitchell et al. 1999a; Dragon et al. 2000; Pogacic et al. 2000). Much of what is known about these proteins comes from their initial characterisation as H/ACA RNP components in yeast via the homologues Cbf5p, Nop10p, Nhp2p and Gar1p. Here it was shown that only Cbf5p, Nop10p and Nhp2p are important for RNP stability and H/ACA RNA accumulation, including exogenously expressed human TERC, but all four proteins are vital for pseudouridylation activity with Cbf5p acting as the catalytic component (Girard et al. 1992; Bousquet-Antonelli et al. 1997; Henras et al. 1998; Lafontaine et al. 1998; Watkins et al. 1998; Dez et al. 2001; Henras et al. 2004b). Subsequent studies have corroborated these findings in mice (Wang et al. 2002; Wang and Meier 2004) and the conserved homology of these proteins across yeast, mouse and humans provided confidence that this would also hold true in the human context, although it had not been confirmed when this study began. Yeast telomerase differs from that of mammals however and is not an H/ACA RNP incorporating these proteins. Consequently characterisation did not extend to their implications in telomerase function and with the exception of dyskerin little is known about their importance in this specialised complex. In order to better understand the role of these core proteins in the molecular pathology of DC, the aim in this part of the project was to investigate their significance for the integrity and function of telomerase.

With catalytic activity guided by the associated H/ACA RNA logic would suggest that activity is dependent on their stable recruitment and accumulation. In human cells the reduction of dyskerin has been associated with a reduction in TERC levels and telomerase activity as well as pseudouridylation (Montanaro et al. 2006). In patients, however, *DKC1* mutations have been associated with reduced TERC levels and telomerase activity (shown *in vitro* to be a direct consequence of the TERC reduction) but not other H/ACA RNAs or pseudouridylation (Mitchell et al. 1999b; Vulliamy et

al. 2001a; Wong and Collins 2006). In contrast, studies in mice have shown that *DKC1* mutations reduce the levels of multiple H/ACA RNAs and impair both pseudouridylation and telomerase activity (Ruggero et al. 2003; Mochizuki et al. 2004).

Knockdown experiments were conducted to systematically target dyskerin, NOP10, NHP2 and GAR1 so that their individual significance in the accumulation of TERC and telomerase activity could be determined. Also, with a view to ratifying the findings of yeast and mouse studies for the first time in human cells, the levels of other H/ACA RNA's were also assessed. Unfortunately pseudouridylation could not be measured directly as the assay (described in Mochizuki et al. 2004) relies on the enzyme RNase T2 to completely digest ribosomal RNA, which has become commercially unavailable.

The discovery of RNA interference (RNAi) profoundly altered the understanding of post transcriptional gene regulation and revolutionised the approach to gene silencing through manipulation of the RNAi pathway (figure 3.1) (reviewed Bantounas et al. 2004; Leung and Whittaker 2005). Now a highly developed tool, RNAi provides simple and effective ways to suppress gene expression levels that are used routinely in the modern lab. As an ideal technology to employ in this study, commercially available short interfering RNAs (siRNAs) pre-designed to target *DKC1* (dyskerin), *NOP10* (NOP10), *NHP2* (NHP2) and *GAR1* (GAR1) transcripts were tested to assess their silencing potential. Successfully efficient candidates were subsequently used in transient knockdown experiments from which functional impacts could be assessed.

A particular concern in siRNA experiments is the question of specificity with several reports of widespread off target effects despite sequence specific targeting (Jackson et al. 2003; Sledz et al. 2003; Lin et al. 2005; Birmingham et al. 2006; Jackson et al. 2006). Falling into two categories, off target effects can be sequence related and specific to each siRNA or sequence independent and caused by siRNAs in general. The introduction of siRNAs (i.e. double stranded RNA - dsRNA) into the cell can mimic a viral threat and trigger an interferon response with widespread consequences on cellular processes such as protein synthesis, cell cycle and apoptosis. Independent of the siRNAs sequence or its target, this is simply a response to the presence of

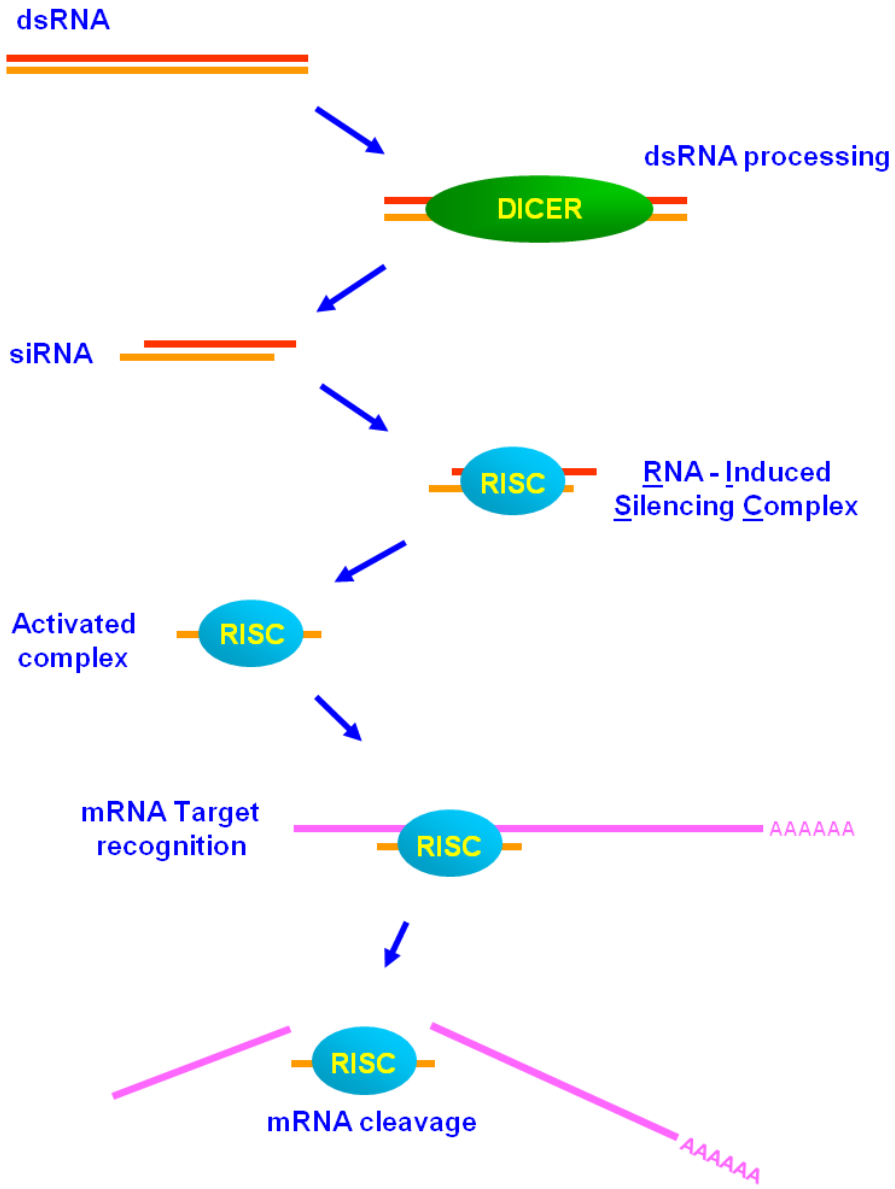


Figure 3.1. The RNAi pathway. Long dsRNA, either endogenous or introduced into the cell, is processed by the Dicer protein into smaller pieces called siRNA. The siRNA is then integrated into the RNA induced silencing complex (RISC). Specific base pairing between the siRNA and mRNA directs targeted mRNA cleavage by the RISC complex and subsequent gene silencing.

siRNA in general and correlates with the amount of the siRNA used. This response is not always activated, however and avoiding high concentrations of siRNA is an effective control against it (Sledz et al. 2003, reviewed Jackson and Linsley 2004). More commonly off target affects are sequence dependant and specific to each siRNA; as with PCR primers, non-specific binding is possible. Complementarity to a 7nt seed region at the 5' end of the siRNA (between base positions 2-8) is not only crucial but also sufficient to elicit an RNAi response (Lin et al. 2005; Birmingham et al. 2006; Jackson et al. 2006). Further, these imperfect matches do not always cause transcript cleavage but may inhibit translation as a method of suppression instead, behaving much like miRNAs (Zeng et al. 2003). Sequence homology to a 7nt siRNA seed region elsewhere in the genome is pretty much unavoidable making sequence specific off target effects almost inevitable and difficult to predict. However, these off target effects are generally less efficient and much weaker by comparison to the targeted effect on the intended transcript (Jackson and Linsley 2004).

Despite the varying implications of off targeting it does not render the technology useless as long as they are considered and controlled for. In the knockdown experiments presented here, possible non-specific effects as a result of the procedure or the addition of siRNA in general are controlled for using mock and negative siRNA transfections. Similarly, sequence specific off targeting is controlled for by using two different siRNAs against each target. Obtaining the same experimental results with independent siRNAs provides support for a target specific effect.

3.2. siRNA Efficiency Testing

3.2.1. Experimental Procedure

Transient transfection experiments were used to determine siRNA silencing efficiencies. Pre-designed siRNAs were obtained in sets of three targeting the same mRNA and a negative control siRNA having no sequence similarity to human gene products (Ambion; detailed in appendix A).

Figure 3.2 provides an overview of the experimental workflow. HeLa cells were transfected with siRNA at an optimal concentration, 30nM for control siRNAs as recommended by the manufacturer, and cultured for 24hr. RNA was extracted, first strand cDNA synthesized and tested for quality and DNA contamination by PCR of the *ABL* gene (primers ABL1 and ABL2 - appendix B; data not shown). Quantitative real time PCR (qPCR) employing TaqMan® probes (details in appendix B) was then used to assess target silencing, final gene expression calculated relative to the untransfected control for comparison (described in chapter 2.2.9).

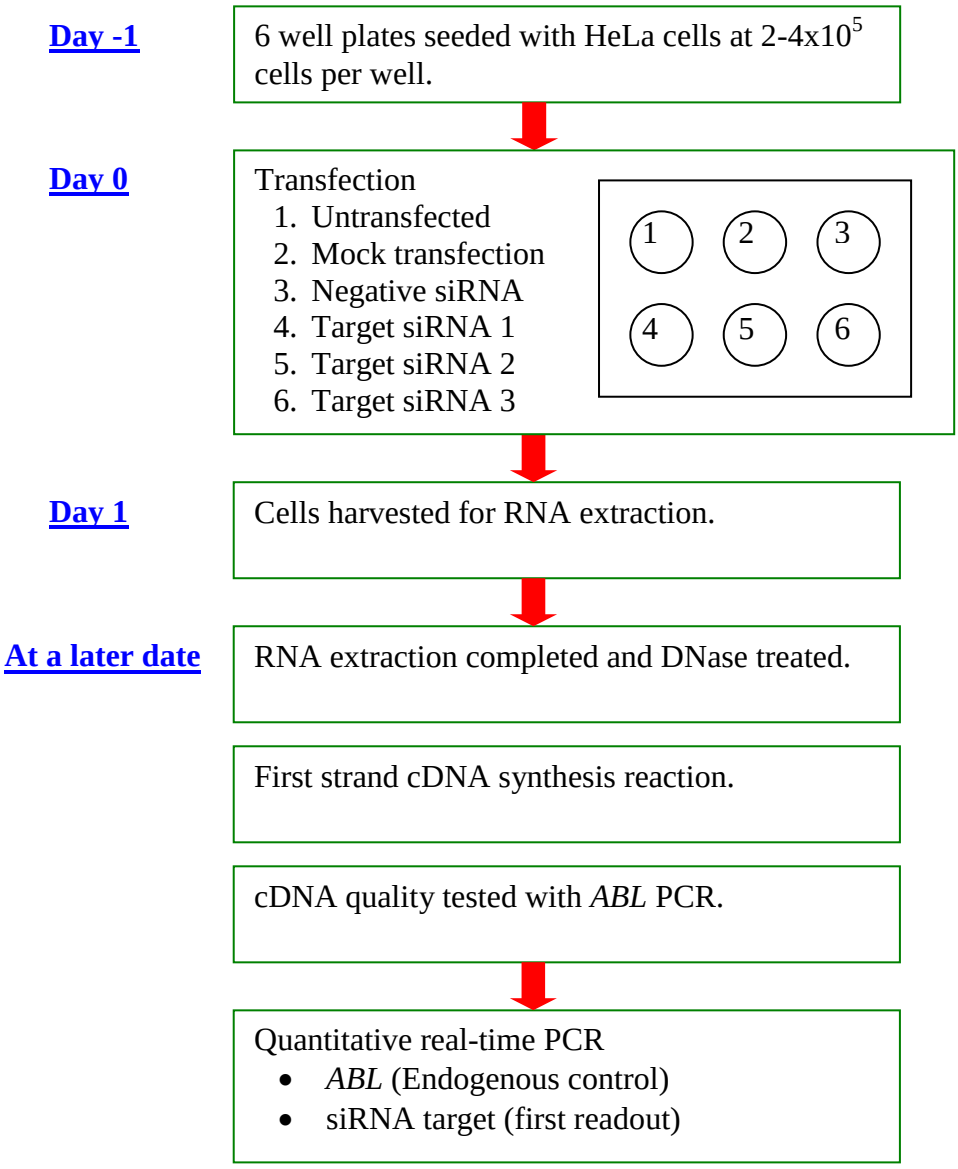


Figure 3.2. siRNA screening – experimental workflow.

3.2.2. Silencing Efficiency Results

Following the manufacturer recommendations for siRNA transfection (as outlined in 2.2.2.1) very little optimization was required. Cell density and confluence at the time of transfection is important for transfection efficiency and cell viability during the experiment. Seeding 2-5x10^5 cells per well was found to result in a suitable confluence for transfection after 24hr with good cell viability. The next step was to optimize siRNA concentrations for a maximal silencing effect. Ambion recommend concentrations between 10-50nM for most cell types, extending to 100nM if necessary and suggesting 30nM as a starting point. Using the *NOP10* siRNAs 30, 50 and 100nM concentrations were tested. Results showed highly efficient silencing (95%+) at all concentrations (figure 3.3). Optimal siRNA concentrations will vary for different targets, however, depending on expression levels and cell type. For reasons of practicality, with regards to sample number and cost, it was decided to screen siRNAs at a single concentration and optimize further only if necessary. Therefore all siRNAs were screened at the upper limit of the recommended range: 50nM.

NOP10 - siRNA 1 was the least efficient and decreased expression by only 70.8% at 100nM. At only 50nM, siRNAs 2 and 3 were more effective and suppressed expression by 96.7 % and 93.5% respectively (figure 3.3).

NHP2 – The most effective siRNAs where 1 and 3, reducing expression by 93.7% and 94.2% respectively. In contrast, siRNA 2 elicited only a 35.9% reduction (figure 3.4a).

GAR1 – Less efficient by comparison, siRNAs 2 and 3 suppressed expression by 79% and 74.8% respectively, siRNA 1 by only 30% (figure 3.4b).

DKC1 - Two sets of siRNAs had to be screened in order to obtain two that were effective. In screen 1 only siRNA 2 was of any value, suppressing expression by 91.4% while siRNAs 1 and 3 only eliciting decreases of 40% and 22.3% respectively (figure 3.4c). In screen 2, siRNAs, 4, 5 and 6 were all highly effective and suppressed expression by 91.9%, 93.8% and 96.6%, respectively (figure 3.4d).

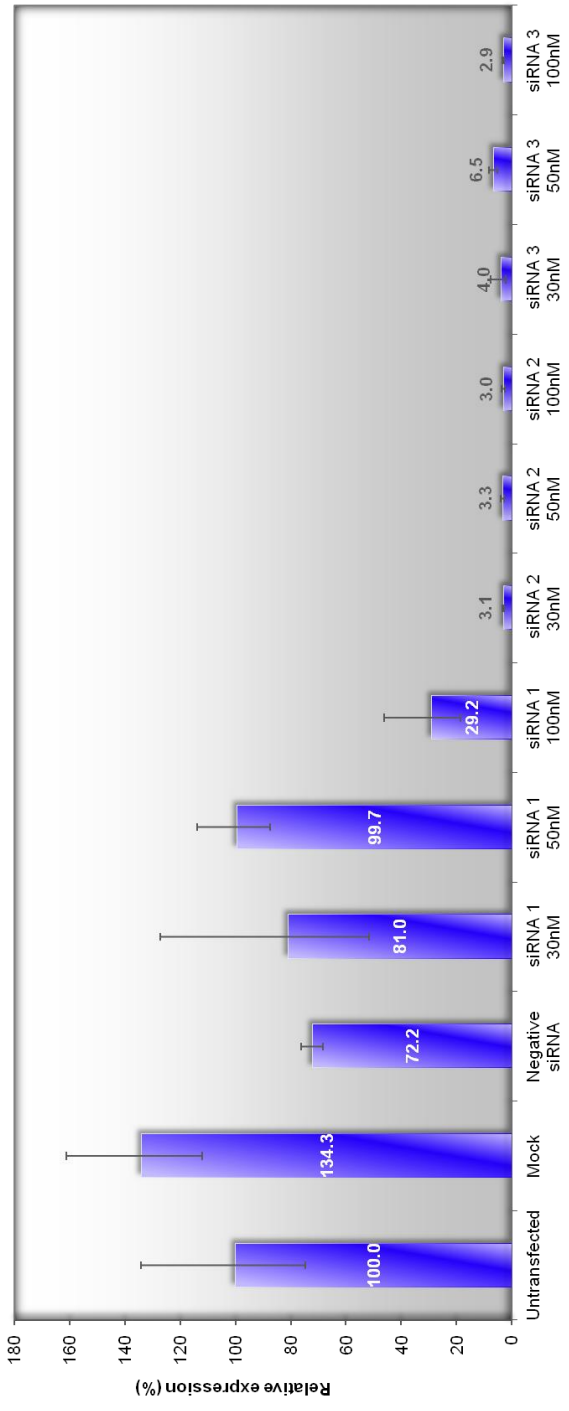


Figure 3.3. *NOP10* suppression by *NOP10* siRNAs. *NOP10* expression levels in HeLa cells 24 hours post transfection with 3 different *NOP10* siRNAs at 30, 50 and 100nM, calculated relative to the untransfected control. Values are from a single experiment, error bars marking the SD between assay triplicates.

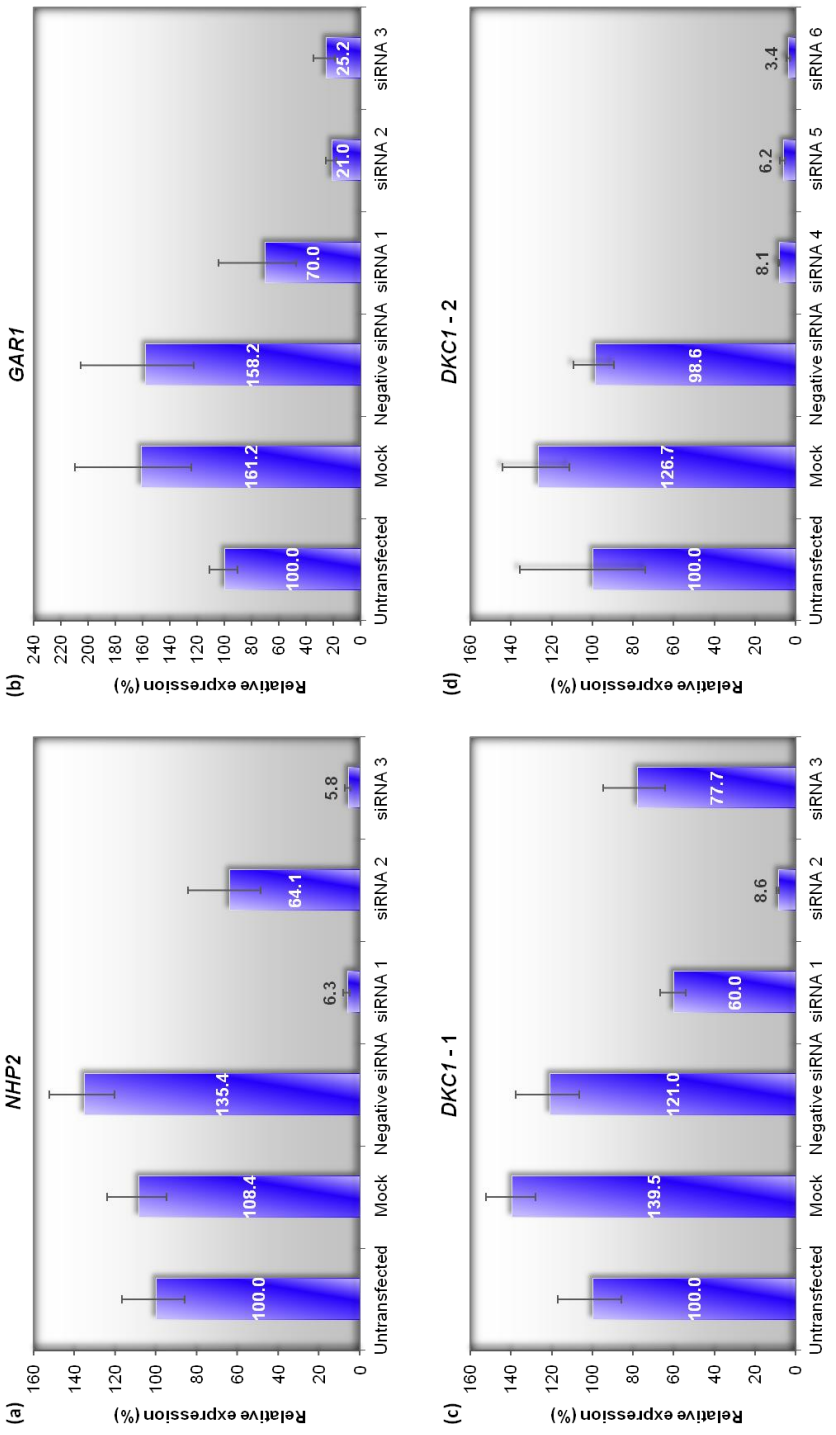


Figure 3.4. Target gene suppression by *DKC1*, *NHP2* and *GAR1* siRNAs. Expression levels of specified targets in HeLa cells 24 hours post transfection with 3 different siRNAs against each, calculated relative to the untransfected control. Values are from single experiments, error bars marking the SD between assay triplicates.

3.2.3. Discussion

The aim here was to test and identify at least two efficient siRNAs to each of the targets: *DKC1*, *NOP10*, *NHP2* and *GAR1*. At a final concentration of 50nM in culture, two out of the three siRNAs tested for *NOP10*, *NHP2* and *GAR1* suppressed gene expression with substantial efficiency. For each siRNA, *NOP10* and *NHP2* expression levels were reduced to less than 10% of that seen in experimental controls and for *GAR1* each siRNA reduced expression to less than 25%. The *DKC1* siRNAs proved to be more problematic and required a total of six (two sets of three) to be screened before the desired target of at least two effective siRNAs was achieved. In fact, four of the six siRNAs tested reduced dyskerin expression to less than 10% compared to experimental controls.

In conclusion, initial siRNA tests were successful in identifying at least two highly effective gene silencers for *DKC1*, *NOP10*, *NHP2* and *GAR1* with the impressive decreases observed suggesting an extremely efficient transfection process. Most importantly, results from two independent siRNAs for each target corroborate to indicate a specific targeted affect which is further supported by the potency with which they suppress gene transcript levels.

3.3. Functional Analysis of Target Gene Knockdown

3.3.1. Experimental Procedure

To determine any functional impacts, the siRNAs identified in chapter 3.2 were used to independently to suppress gene expression and knockdown the levels of dyskerin (*DKC1*), NOP10 (*NOP10*), NHP2 (*NHP2*) and GAR1 (*GAR1*) in HeLa cells. Extending the time course of the experiment to allow downstream effects to be seen, the same transient transfection system described in chapter 3.2.1 was implemented with samples collected at 48 and 72hr post transfection for analysis. For each target both siRNAs were used independently at final concentrations of 50nM, figure 3.5 provides a detailed overview of the experimental workflow. Knockdown experiments were conducted in duplicate and functional data averaged to provide the final result.

Note: - Technical difficulties in some stages required the NOP10 and NHP2 experiments to be performed a third time in order to obtain all the functional data. Target gene silencing and TERC levels represents the average of all three experiments, TRAP activity and H/ACA RNA levels the average of only two.

3.3.1.1. Target Gene Silencing

Target gene silencing was assessed at 48 and 72hr. As before, RNA was extracted and first strand cDNA synthesised. DNA contamination and cDNA quality were assessed by *ABL* PCR (primers ABL1 and ABL2 - appendix B; data not shown). Target gene expression was then determined by qPCR using TaqMan® probes (details in appendix B), calculated relative to the 48hr untransfected control for comparison (described in chapter 2.2.9).

In siRNA experiments it is beneficial to validate target knockdown at the protein level. However, suitable antibodies were not readily available for all of the targets, only dyskerin courtesy of an antibody generously gifted by Yves Henry. Western blotting (described in chapter 2.2.25) was used to verify dyskerin levels against those of a β-actin loading control at 48hr post transfection. For each sample, 30µg of protein were loaded onto a 12% gel and separated under denaturing and reducing conditions. Immunoblotting was then performed using rabbit anti dyskerin (Yves Henry) and mouse anti β-actin (Abcam) primary antibodies in combination with goat anti rabbit-HRP (BD) and goat anti mouse-HRP (BD) secondary antibodies.

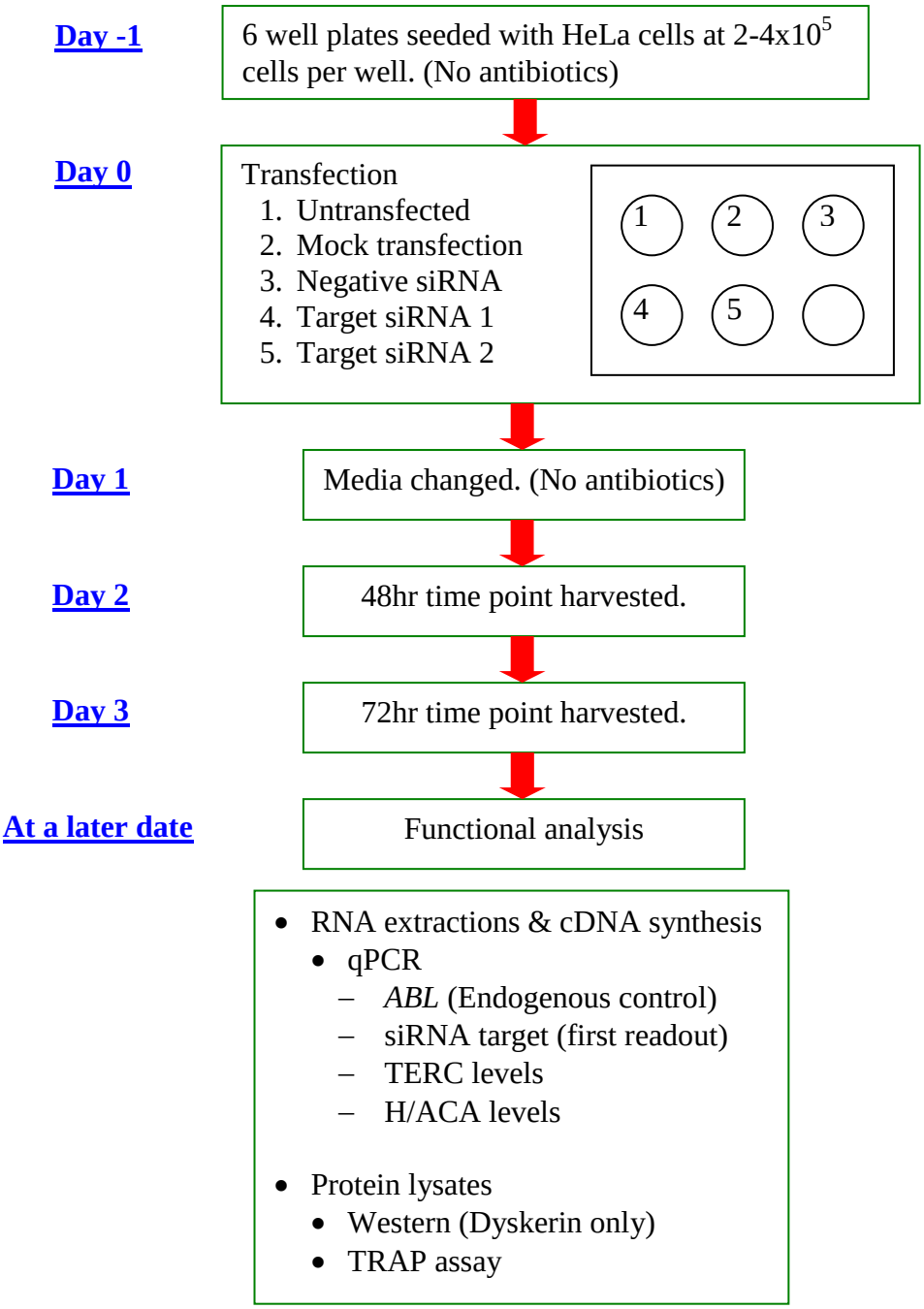


Figure 3.5. Transient siRNA knockdown – experimental workflow.

3.3.1.2. TERC Accumulation

Impacts on TERC accumulation were assessed at 48 and 72hr. TERC levels were determined by PCR using a TaqMan® probe (details in appendix B), calculated relative to the 48hr untransfected control for comparison (described in chapter 2.2.9).

3.3.1.3. TRAP Activity

Functional impacts on telomerase activity were assessed at 48hr post transfection. Extracts were prepared from cells collected at this time point and telomerase activity measured using the TRAP assay (described in chapter 2.2.24), activity levels calculated relative to the untransfected control for comparison.

3.3.1.4. H/ACA RNA Accumulation

To investigate any affects on H/ACA RNAs more generally a set of 5 different RNAs from this class: E3, U64, U68, U70, U108, were quantified alongside TERC and the control C/D box RNA U3. Preparing a standard curve for each, absolute copy numbers were determined in the cDNA samples by quantitative real time PCR using SYBR green (described in chapter 2.2.9, primers in appendix B). Measured 48 and 72hr post transfection, values are calculated as a target/*ABL* ratio and expressed relative to the untransfected controls for comparison.

3.3.2. Results

3.3.2.1. Target Silencing Efficiency

Corroborating with results from the siRNA efficiency tests, when used at 50nM the siRNAs against *DKC1, NOP10, NHP2* and *GAR1* all successfully reduced target gene expression compared to the experimental controls (figure 3.6). In each case the decreases were sustained over the 72hr with comparable results for both siRNAs and experimental duplicates. In summary, both siRNAs against *NOP10* and *NHP2* reduced expression to less than 8% and 6% respectively (figures 3.6a & b), *GAR1* siRNAs reducing expression to below 30% (figure 3.6c). The *DKC1* siRNAs performed less efficiently than in the initial tests with *DKC1* expression remaining as high as 36.9–40.8% for siRNA 4 and 17.7–32.6% for the siRNA 5 (figure 3.6d). However, this was still a substantial decrease.

Western blotting performed on samples taken from the dyskerin experiment at 48hr successfully identified bands corresponding to dyskerin (expected size ~57kDa) and β-actin (expected size ~42kDa) (figure 3.7). Measured across both the experimental duplicates, the data clearly shows that dyskerin protein levels were substantially reduced in the siRNA treated samples compared to the experimental controls.

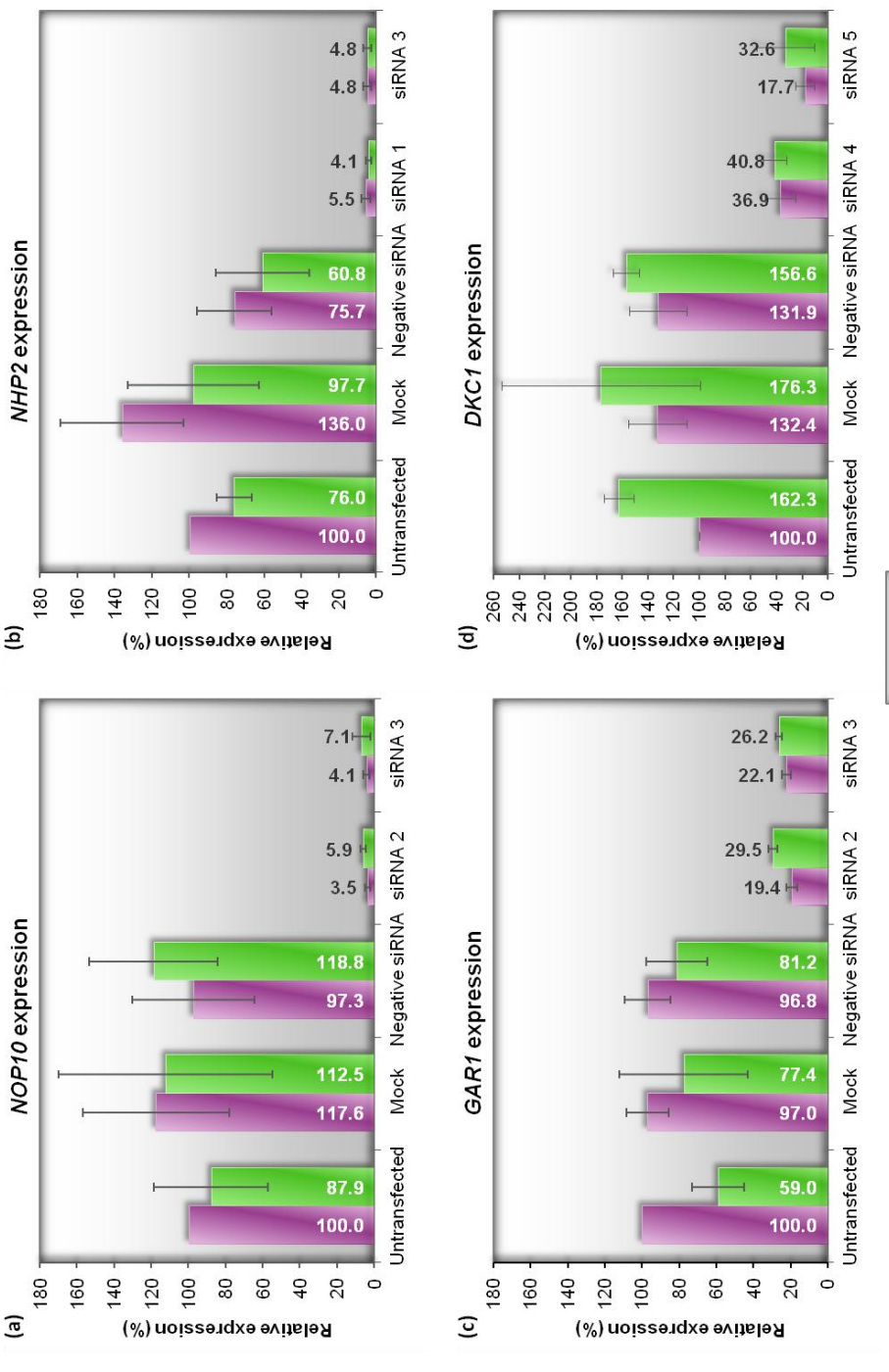


Figure 3.6. Target gene expression in the transient knockdown experiments. Measured in HeLa cells at 48 and 72 hours post transfection with siRNAs against each target and calculated relative to the 24 hour untransfected control. Values are the average of duplicate (c & d) or triplicate (a & b) experiments with error bars marking the SD.

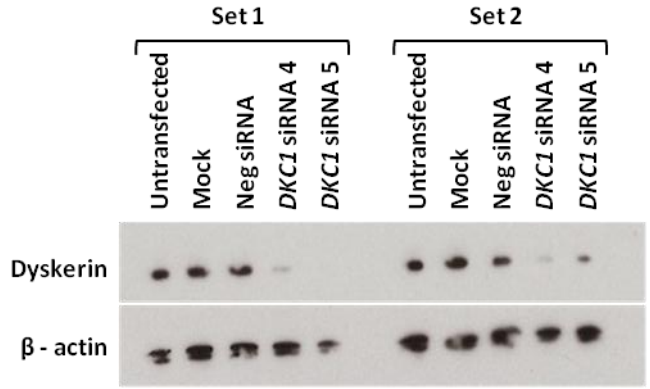


Figure 3.7. Western blot data showing dyskerin knockdown. Dyskerin levels were analysed in samples from both duplicates of the dyskerin knockdown experiment as indicated. β-actin levels were measured as a loading control.

3.3.2.2. TERC Levels

TERC levels were drastically reduced from 48hr onward in response to the knockdown of either dyskerin, NOP10 or NHP2 but not GAR1 (figure 3.8). In the dyskerin experiment this effect was slightly staggered over the time course, TERC levels falling to between 35-45% by 48hr before reaching a low of 18-25% at 72hr (figure 3.8d). A similar effect was also evident in the NHP2 experiment but not with both siRNAs. TERC levels fell to 37.8% by 48hr and then 15.2% at 72hr with siRNA 3, but for siRNA 1 TERC levels were between 15-20% at both time points (figure 3.8b). Knocking down NOP10 reduced TERC levels to between 20-25% at both time points (figure 3.8a). Despite these subtle variations, by 72hr TERC levels are similarly reduced to between 15-25% in each of these three knockdown experiments. In contrast, TERC levels under the suppression of GAR1 reveal only inconsistent fluctuations, similar to those seen in the controls, and remain consistently higher at 60-100% (figure 3.8c). Unfortunately, due to the limited number of experimental replicates, meaningful statistical analysis of these results was not feasible.

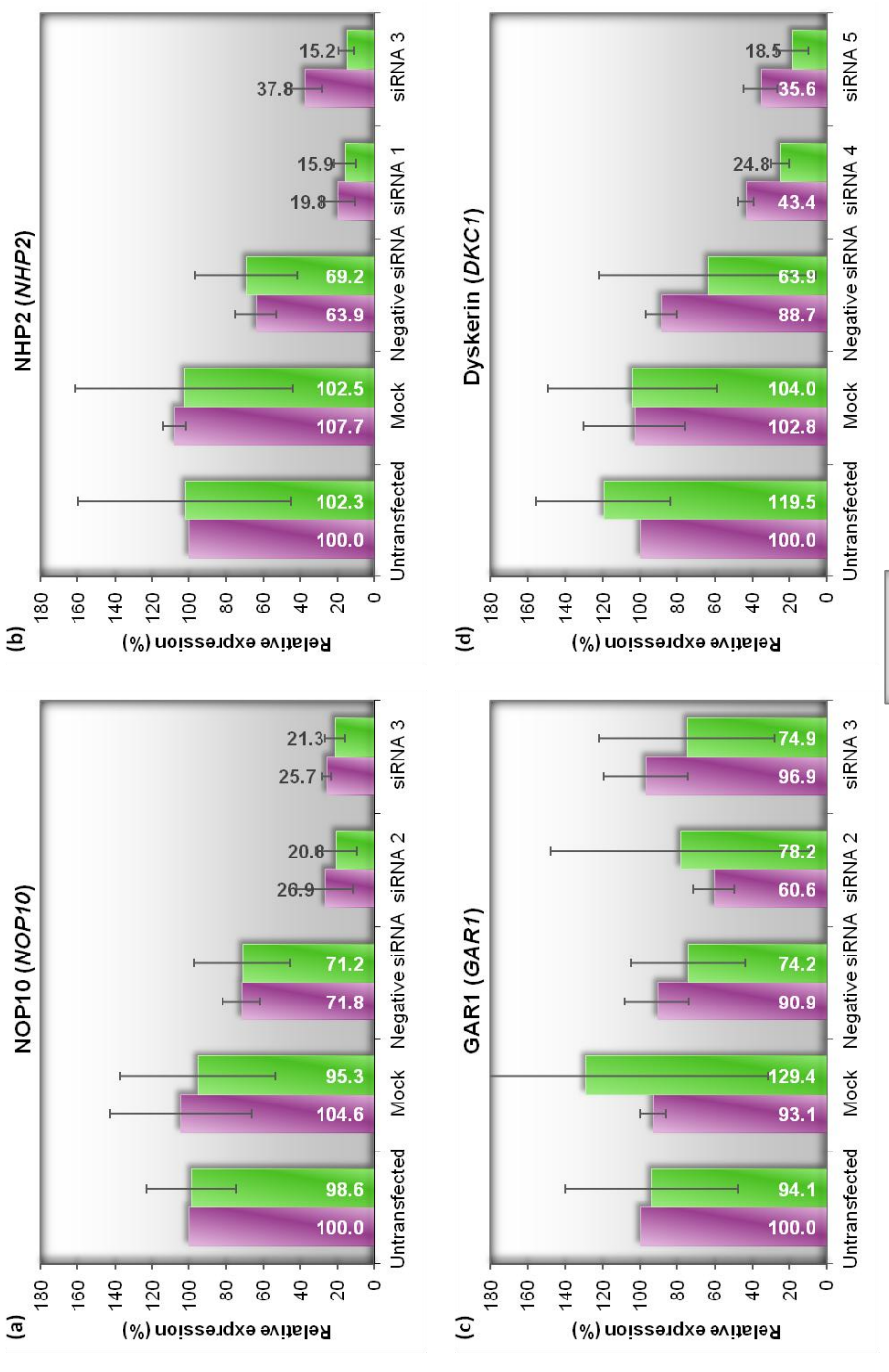


Figure 3.8. TERC levels in the transient knockdown experiments. Measured in HeLa cells 48 and 72 hours post transfection with siRNAs against each specified target and calculated relative to the 24 hour untransfected control. Values are the average of duplicate (c & d) or triplicate (a & b) experiments with error bars marking the SD.

3.3.2.3. Telomerase Activity Levels

TRAP assay analysis at 48hr indicates that telomerase activity was severely impaired by the knockdown of dyskerin or NOP10 but not NHP2 or GAR1 (figure 3.9a-d). The most dramatic impact was seen in the NOP10 experiment where activity dropped to below 28% compared to controls with both siRNAs (figure 3.9a). Less strikingly, in the dyskerin experiment activity was reduced to below 47% for one siRNA and 29% for the second (figure 3.9d). Conversely, telomerase activity appeared largely unaffected in both the GAR1 and the NHP2 experiments where in each case activity was sustained above 80% compared to controls (figures 3.9b & c). Once again, meaningful statistical analysis of these results was precluded by the small number of experimental replicates.

Unfortunately, in each knockdown experiment the negative siRNA also reduced telomerase activity to between 23-70%, suggesting a non-specific off target influence that potentially undermines these results. Off target affects in RNAi experiments are discussed in chapter 3.1. Sustained activity in the GAR1 and NHP2 experiments excludes the simple presence of siRNA as the negative influence, which is what this control would indicate. Instead, this signifies something specific to the negative siRNA used and suggests a sequence dependent off target effect. To investigate this further the negative siRNA was compared against three other commercial alternatives for impacts on this assay. HeLa cells were transfected with the four different negative siRNAs (details in appendix A), a *GAPDH* siRNA and *NOP10* siRNA 2 (each at 50nM) alongside mock and untransfected controls. TRAP assay analysis was then performed on samples taken after 48hr (figure 3.10). As a positive control for an effect on telomerase activity the *NOP10* siRNA dramatically reduced TRAP activity, reproducing the results of the NOP10 experiment. Acting as a ‘mock’ negative control for an effect on telomerase activity (positive siRNA silencing of a target having no association with telomerase) the *GAPDH* siRNA did not affect TRAP activity as expected. In contrast, all four negative siRNAs variably impacted on telomerase activity. Indicating that this effect is not specific to one negative siRNA, this would suggest a sequence independent off target effect. Once again though, this is contradicted by the other controls which demonstrate that the transfection procedure itself or the use of siRNAs generally does not have a negative impact.

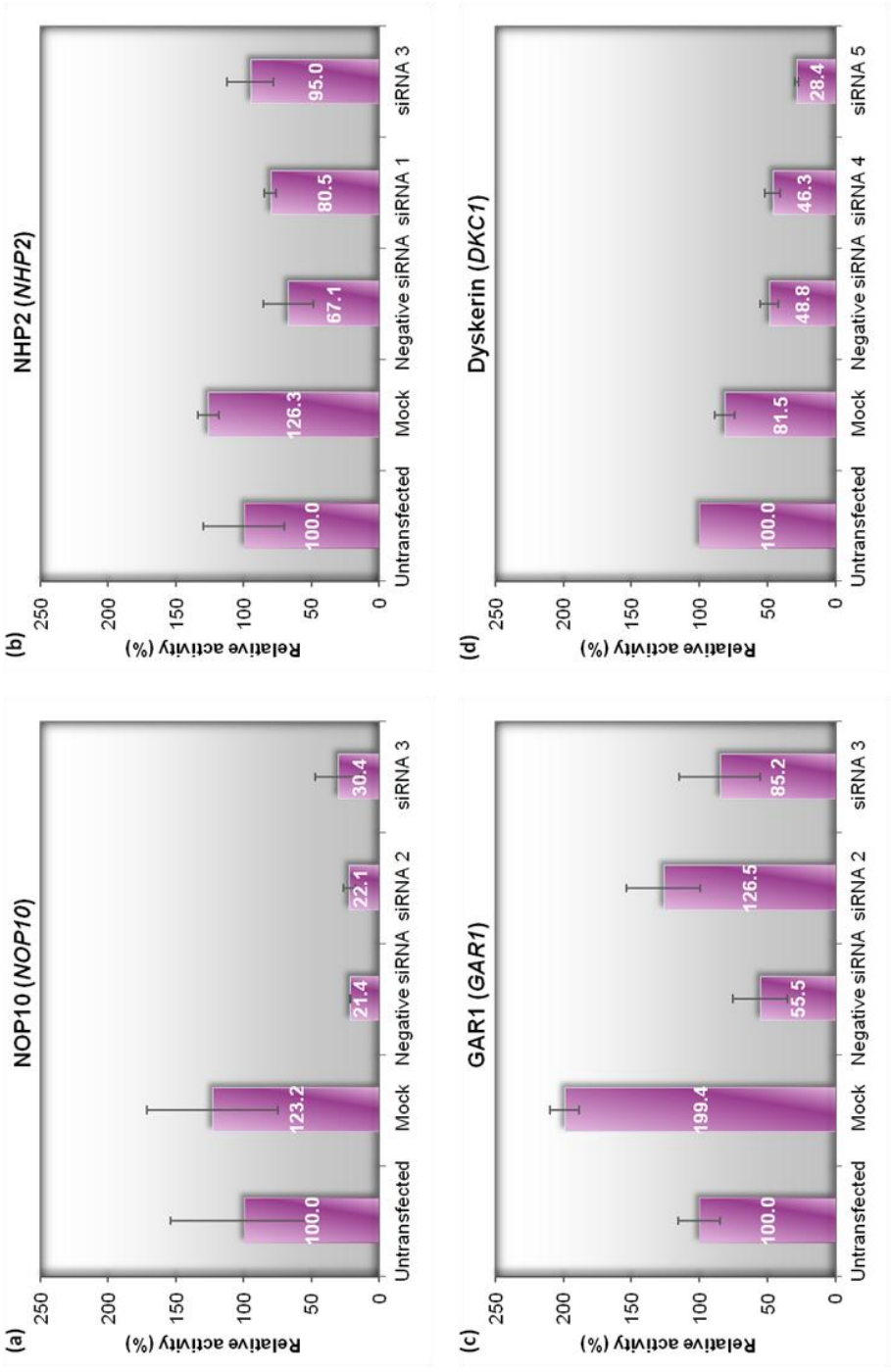


Figure 3.9. TRAP activity in the transient knockdown experiments. Measured in HeLa cells 48 hours post transfection with siRNAs against each specified target and calculated relative to the untransfected control. Values are the average of duplicate experiments with error bars marking the SD.

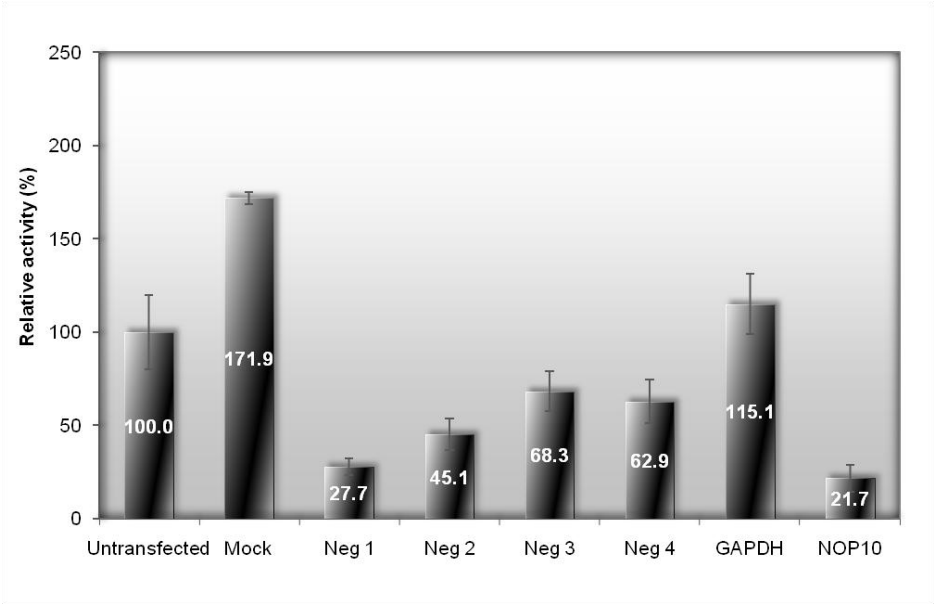


Figure 3.10. TRAP activity in HeLa cells transfected with negative siRNA controls. Measured at 48 hours post transfection and calculated relative to the untransfected control, values are the average of duplicate experiments with error bars marking the SD. (Neg1 corresponds to that used in the transient experiments)

3.3.2.4. H/ACA RNA Levels

Striking fluctuations were observed when quantifying the H/ACA RNAs: TERC, E3, U64, U68, U70, U108, and the C/D box RNA U3, in these knockdown experiments (figure 3.11). Inconsistency between controls and experimental duplicates was evident across the different experiments, compounded by further variability from 48 to 72hr. This was a particular problem in the GAR1 experiment which has subsequently been excluded here on this basis. Results for the siRNA treated samples were generally more consistent, though in some cases disagreement between the two siRNAs was observed. It is difficult to discern if this variability is the result of biological fluctuations, experimental influences or a combination of the two. Therefore, it is difficult draw any firm conclusions from this data. However, it does provide some indication of how the H/ACA RNAs may have been affected in these experiments.

One of the more consistent results, TERC levels agree with those of the previous analysis and reiterate a dramatic reduction in response to the knockdown of dyskerin, NOP10 or NHP2. Also, again there is suggestion of a gradual effect in which TERC

levels fell considerably by 48hr (figure 3.11a, c, & d) but did not reach their lowest until 72hr (figure 3.11b & e). Expanding the analysis to the other H/ACA RNAs measured shows indications of a variable affect that differed depending on the knockdown target.

Reducing NHP2 appeared to have the broadest potential impact with both siRNAs exhibiting substantial and comparable decreases in all of the H/ACA RNA tested at both 48 and 72hr (figure 3.11d & e).

In the NOP10 knockdown there were indications of a decrease in U68, U70 and U64 at both 48 and 72hr with a possible, more subtle decline in U108 as well (figures 3.11a & b).

Surprisingly, knocking down dyskerin appeared to have less impact, possibly reflecting the reduced efficiency of the siRNAs in this experiment. After 48hr only U68 was found to be substantially reduced, though E3, U64 and U70 were also suggested to suffer small decreases (figure 3.11c). The 72hr data has been excluded here because of the issues discussed earlier but it is worth noting that the results would indicate more substantial decreases by that time.

Finally, in all of the knockdown experiments the U3 C/D box RNA control proved to be erratic, correlating poorly between experimental duplicates and time points. However, best indications from the averaged data are that this RNA was not negatively affected by the suppression of either dyskerin, NOP10 or NHP2 (figure 3.11a-f)

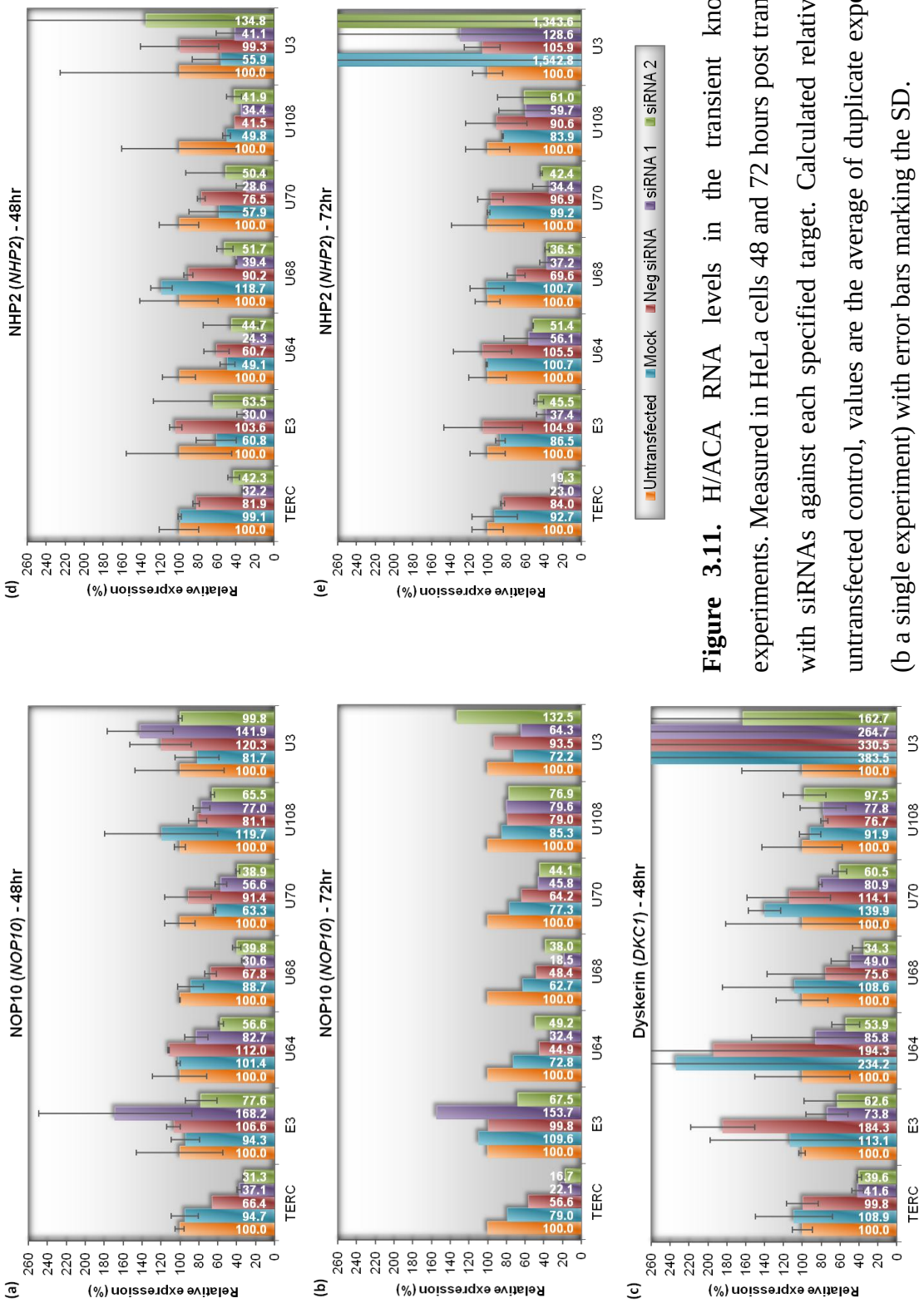


Figure 3.11. H/ACA RNA levels in the transient knockdown experiments. Measured in HeLa cells 48 and 72 hours post transfection with siRNAs against each specified target. Calculated relative to the untransfected control, values are the average of duplicate experiments (b a single experiment) with error bars marking the SD.

3.3.3. Discussion

Aiming to effect independent transient knockdown experiments against dyskerin, NOP10, NHP2 and GAR1, siRNAs were successfully used to reduce their gene expression over a 72hr period. Reflecting results from the initial efficiency tests, *NOP10* and *NHP2* suppression was highly effective (90%+ reductions) along with *GAR1* (70%+). *DKC1* suppression (60%+) did not meet expectations following the initial tests, possibly indicating reduced transfection efficiency, however, there was still a considerable reduction in expression with supporting data to confirm a substantial decrease in dyskerin protein levels by 48hr.

Studies in yeast and mice have demonstrated that the orthologues of dyskerin, NOP10 and NHP2 but not GAR1 are important for the stable accumulation of H/ACA RNAs including TERC. Seeking to investigate this for the first time in human cells (with the exception of dyskerin), the results here confirm that TERC levels were dramatically reduced as a direct consequence of knocking down either dyskerin, NOP10 or NHP2 but not GAR1 (published Walne et al. 2007; Vulliamy et al. 2008). Expanding this analysis to a panel of other H/ACA RNAs posed some issues but suggests there were some decreases and indicating a broader negative impact on these RNAs in response to knocking down dyskerin, NOP10 or NHP2. This impact is variable and does not affect all the RNAs equally, also differing depending on the protein targeted. In the yeast studies subtle variations were similarly observed with some RNAs affected more severely than others (Dez et al. 2001). One possible explanation could be that it reflects variation in their individual expression levels and stability. With regards to this analysis in the GAR1 knockdown experiment, poor consistency between experimental duplicates led to the data being excluded, however, best indications would suggest there was little or no affect on H/ACA RNA levels (data not shown).

Within these data sets there are suggestions that the impact on RNA levels was slightly staggered over the 72hr period, particularly in the dyskerin experiment. It is probable that this reflects a staggered decrease of the target protein. In knockdown experiments such as these it is important to consider the issue of protein stability and half life. This determines how quickly transcript suppression translates through to the protein level and directly influences how long the effects of the knockdown take to manifest. Therefore, while target expression levels indicate a high level of transcript

suppression from 24hr onward a delayed or gradual effect at the protein level is possible. Unfortunately, protein levels could only be verified in the dyskerin experiment. With protein levels only analysed at 48hr, a time staggered effect was not investigated but instead served to confirm that dyskerin protein levels were substantially suppressed in the siRNA treated samples by this time. The potential for a delayed effect should be considered in the GAR1 experiment where protein levels could not be verified. However, the studies in yeast demonstrated depletion of the Gar1p protein within 12hr of its suppression as well as the absence of any impact on H/ACA RNA levels within 48-72hr (Girard et al. 1992; Dez et al. 2001).

Together, these results agree with those of the yeast and mice studies and indicate that dyskerin, NOP10 and NHP2 but not GAR1 are important for the stability and accumulation of H/ACA RNAs in human cells. Studies in mice have indicated that GAR1 is not involved in H/ACA RNA recruitment by demonstrating that dyskerin, NOP10 and NHP2 form a core trimer which binds the H/ACA RNA before associating with GAR1 (Wang and Meier 2004; Darzacq et al. 2006). This is supported by the crystal structure of a complete archaeal H/ACA RNP complex which has been published during the course of this study (figure 3.12) (Li and Ye 2006). In a triangularly shaped complex, NOP10 (Nop10) is tightly sandwiched between dyskerin (Cbf5) and NHP2 (L7Ae), which form two of the corners. As if stretched between these corners the H/ACA RNA component reaches across and interacts with all three of these proteins. Situated away from this RNA binding, GAR1 makes up the third corner of the complex and interacts only with dyskerin. In this structure the core trimer and its importance for H/ACA RNA binding and recruitment is clearly visible while it is apparent that GAR1 is not directly involved in this interaction. Instead GAR1 has been implicated in modulating the conformation of a loop structure identified in dyskerin, the thumb loop, which regulates substrate turnover by controlling the loading and release of the target substrate RNAs (Li and Ye 2006; Li et al. 2011).

Investigating telomerase activity, TRAP assay analysis at 48hr suggests that this was impaired as a result of reducing NOP10 and dyskerin but not NHP2 or GAR1. Studies have shown that depleting dyskerin levels has a detrimental impact on both TERC levels and telomerase activity (Montanaro et al. 2006) and it has been demonstrated

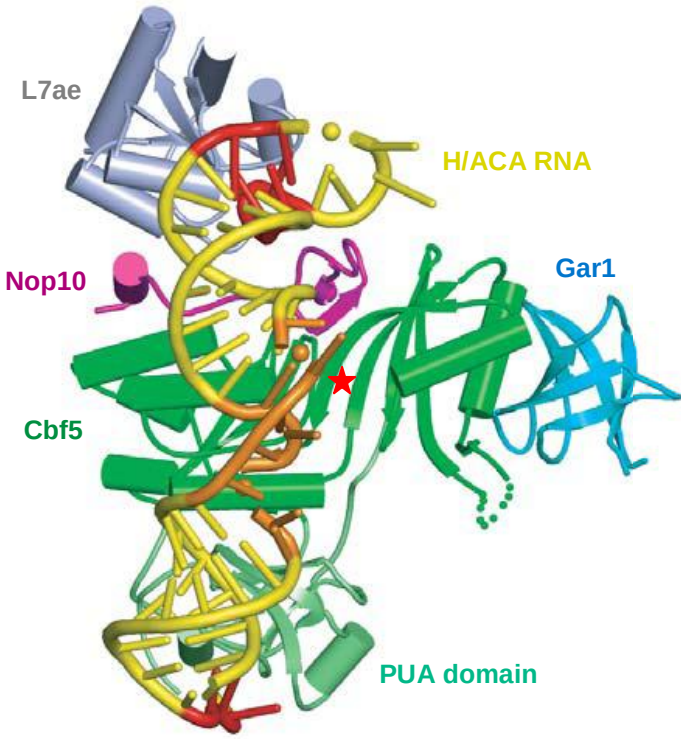


Figure 3.12. The crystal structure of a *Pyrococcus Furiosus* (Pf) H/ACA RNP. Each component is coloured differently, the PUA and catalytic domains of Cbf5 also shaded differently and a red star marking the active site. (Figure adapted from Li and Ye 2006).

that TERC levels are a limiting factor for telomerase activity (Fu and Collins 2003; Marrone et al. 2004; Wong and Collins 2006). Consequently, based on TERC levels in these experiments, the knockdown of dyskerin, NOP10 or NHP2 was expected to adversely affect telomerase activity. Therefore, the lack of a discernible impact in the NHP2 knockdown was somewhat surprising and contradicts the established correlation between TERC levels and telomerase activity. Certainly, similar impacts on TERC levels in the NOP10 and dyskerin experiments had a negative effect. It is difficult to rationalise this result, which is reproducible between siRNAs and experimental duplicates. Yeast studies have shown that H/ACA RNPs readily form in a cell free context, more importantly that incomplete but functional complexes can form in the absence of Nhp2p and retain more than 50% activity (Henras et al. 2004b; Li et al. 2011). However, This does not reflect the impact on H/ACA RNA levels and pseudouridylation when Nhp2p is depleted in yeast which suggests that these

complexes are not stable *in vivo* (Henras et al. 1998; Dez et al. 2001). It is interesting to speculate the possibility of similar complexes forming in the cell lysate used for the TRAP assay, perhaps providing levels of activity that are not a true reflection of that *in vivo*. Further experiments are needed to investigate this possibility.

GAR1 suppression had no influence on TERC levels at 48hr and here the TRAP assay results suggest that it does not play a more direct role in telomerase activity either. Studies have shown that GAR1 associates with the H/ACA RNP post RNA recruitment and speculate that its role is not central to the stability of the complex (Darzacq et al. 2006). The results here would support this, indicating that GAR1 is not implicated in either the stability or catalytic activity of telomerase. A distinction between telomerase activity and function should be made at this point, the catalytic process of transcription being its activity and the extension of telomeres its function. It is therefore important to note that successful telomerase function is as dependant on its correct localisation as it is its catalytic activity and a failure in either would be detrimental to telomere maintenance. The TRAP assay measures telomerase activity in homogenised cell lysates, free of cellular compartmentalisation and any restrictions or regulatory influences this may impose. Consequently factors such as localisation that may affect telomerase function rather than direct activity may be neutralised and undetected in this assay. In consideration of this, while these results indicate that GAR1 is not required for the formation of a stable, catalytically active telomerase complex they do not exclude the possibility that GAR1 may still have some other critical role in telomerase function and telomere maintenance.

Unfortunately the negative siRNA control also had a negative impact that implies off target influences were affecting telomerase activity in these experiments, potentially undermining the validity of these results. The lack of a similar impact using the *GAR1* and *NHP2* siRNAs suggest this effect to be specific to the negative siRNA. To investigate, the negative siRNA was compared against three other commercial alternatives for effects on this assay and surprisingly all four caused a variable reduction in TRAP activity. This result adds complexity rather than clarity to the situation. Four different negative siRNAs causing the same detrimental effect would point toward a sequence independent off target influence but this is countered by normal activity in the mock transfection and *GAPDH* siRNA samples. It is unlikely

that these independent siRNAs generate the same profile of sequence specific off targets, though some overlap could be expected. However, with the potential to affect numerous cellular pathways it is quite feasible that the broader cellular consequences of off targeting could overlap. Indeed, studies have shown that some siRNAs, including non-targeting controls, can induce adverse affects on cell cycle, proliferation and viability. Having confirmed that neither an interferon response or target specific effects were the cause this was attributed to sequence dependent off targeting bringing about changes in gene expression that culminate in a common overall effect (Fedorov et al. 2006; Tschaharganeh et al. 2007). Telomerase activity is a tightly regulated process and is influenced by cell cycle and proliferation (Greider 1998). Conceivably, the effects observed here could be the result of direct changes in telomerase associated pathways or indirectly through consequential changes in cellular processes that influence its activity. Unfortunately, either way this serves to diminish the value of the TRAP assay in these experiments. Importantly, however, this is not a concern when analysing target expression and TERC levels where the results for the negative siRNA generally agree with those of the untransfected and mock transfection controls.

The important question is: are the results of the TRAP assay reliable and informative? Indications are that this effect is not generic to all siRNAs so it is possible that in these experiments the effect is isolated to the negative siRNA control. In the GAR1 and NHP2 experiments there is no cause to doubt this as the knockdown of either failed to elicit an effect on telomerase activity, targeted or other. Instead the difficulty arises when considering the NOP10 and dyskerin knockdown experiments, but there is support for a genuine target mediated effect. In both cases there is consistency between the two siRNAs which reinforces the result in favor of a targeted effect. Also, both knockdowns have an impact on TERC levels that would be expected to impair telomerase activity as observed. Finally, with regards to dyskerin there is already evidence of its importance for telomerase activity in other studies and mouse models presented in the literature (Mitchell et al. 1999b; Mochizuki et al. 2004; Wong et al. 2004; Montanaro et al. 2006; Wong and Collins 2006; Montanaro et al. 2008). Therefore, while acknowledging the limitations imposed by the negative siRNA, in view of supporting evidence elsewhere, the inclination is to accept these results as a

good indication that knocking down either NOP10 or dyskerin has a direct negative impact on telomerase activity.

In conclusion, these experiments show that dyskerin, NOP10 and NHP2 are important for the stable accumulation of TERC and other H/ACA RNAs, demonstrating this for the first time in a human context (except for dyskerin). This would support these proteins having a central role in H/ACA RNP stability including telomerase, but only dyskerin and NOP10 were shown to be important for telomerase activity. Based on its significance for TERC levels, however, NHP2 was expected to be similarly implicated in this activity and further investigation of this is warranted. In contrast GAR1 does not appear to be significant for either TERC levels or telomerase activity in these experiments, though an indirect effect cannot be ruled out.

Chapter 4. Stable Knockdown Experiments

4.1. Introduction

Transient siRNA experiments offer an easy method for determining the immediate effects of suppressing a gene of interest and as such provided a beneficial approach for the functional characterisations described in chapter 3. However, some downstream consequences take longer to manifest than the experiment allows, changes in telomere length being a relevant example. Inversely, some consequences may only be transient and not sustained over time. In view of this it was sought to confirm and clarify the findings of the transient experiments in knockdown studies that could be sustained and monitored over a longer period.

In order to achieve long term knockdown experiments a stable system of gene suppression is required. This could be achieved by adapting the method of RNAi from siRNAs to the more sustainable short hairpin RNA (shRNA) mediated approach. In summary, shRNA is expressed in the cell from a DNA construct encoding the sense and antisense strands of the siRNA separated by a short loop sequence. The RNA transcript folds at the loop and self anneals to form a dsRNA stem loop structure: a short hairpin RNA. The shRNA is structurally related to endogenous miRNAs and is processed accordingly by the RNAi machinery to mediate target silencing (figure 4.1). Conventionally the RNA polymerase III promoters U6 and H1 are used for shRNA expression, generating transcripts lacking a long poly (A) tail that may hinder subsequent RNAi processing machinery. However, RNA polymerase II promoters such as the CMV promoter have been used. (Brummelkamp et al. 2002; Gou et al. 2003; Amarzguioui et al. 2005)

Attempts were made to create shRNA expression vectors for generating stable cell lines that express shRNA against *DKC1*, *NOP10*, *NHP2* and *GAR1*, where the silencing effect could be sustained indefinitely.

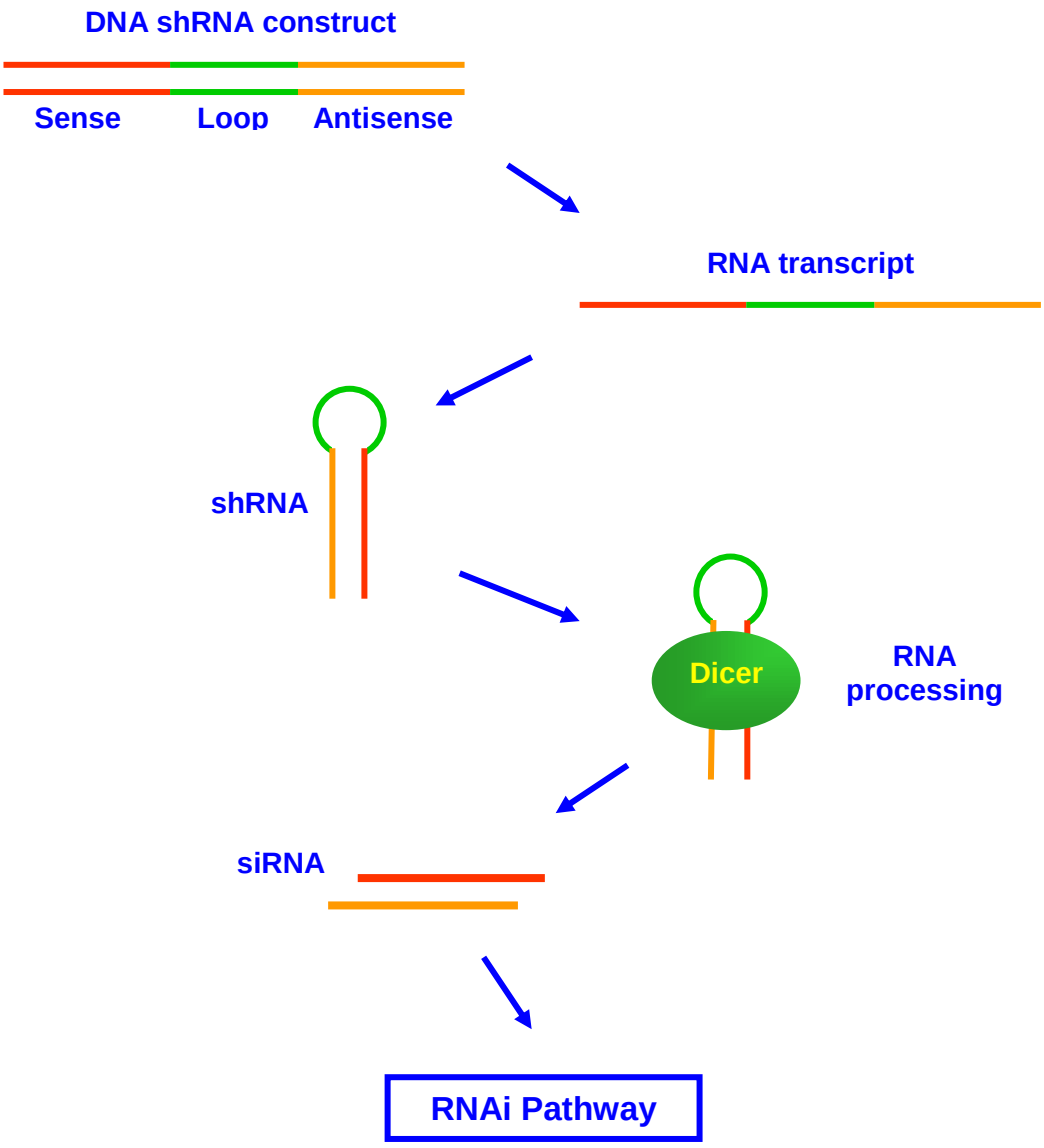


Figure 4.1. shRNA mediated RNAi. A DNA construct encodes the sense and antisense sequences separated by a short loop sequence. The expressed RNA transcript folds at the loop allowing complementary sense and antisense sequences to anneal forming the short hairpin RNA (shRNA) structure. This double stranded RNA molecule is then recognised by the Dicer protein and is processed into siRNA as it enters the RNAi pathway as depicted in figure 3.1.

4.2. PCR Strategy

Initial attempts to create the shRNA expression constructs followed a PCR based strategy described in the literature (Castanotto et al. 2002; Gou et al. 2003). In a two step, overlapping PCR process depicted in figure 4.2, shRNA and U6 promoter sequences are linked to create an shRNA expression cassette. The expression cassettes can be used independently for transient effects or subsequently cloned into selectable plasmid expression vectors for stable transfection experiments.

4.2.1. Experimental Procedure

Sequence data for the human U6 promoter (AY623053) and the siRNAs used in chapter 3 were used to design an shRNA expression cassette against *NOP10* and appropriate oligonucleotides for its contsruction (detailed in appendix B). Using genomic DNA as a template, the human U6 promoter was PCR amplified using the hU6 Forward and hU6 Reverse primers (details in appendix B) in a standard PCR reaction (described in chapter 2.2.8). The hU6 Forward primer is complementary to 22nt in the 5' end of the U6 promoter with a *MluI* linker appended to the 5' end, used in all subsequent PCRs. The hU6 Reverse primer is complementary to 19nt at the 3' end of the U6 promoter and has a *NotI* linker appended to the 5' end. The purified U6 promoter was subsequently used to create the *NOP10* shRNA expression cassette in the two step PCR process summarised in figure 4.2, using the universal hU6 forward primer and the reverse primers *NOP10* shRNA 1 and 2. Standard reaction conditions were used in both steps (described in chapter 2.2.8) and products were purified by gel extraction between steps. To increase product yield, step 2 reactions were scaled up from 25 to 50µl.

PCR 1: Purified U6 promoter acts as template in a PCR with the hU6 Forward and *NOP10* shRNA 1 primers. 5' to 3' the *NOP10* shRNA 1 primer has 7nt of loop sequence, 19nt of sense strand sequence and 9nt of complementary sequence with the amplified U6 promoter 3' end. The result is attachment, and amplification, of the shRNA sense strand and loop sequence to the 3' end of the U6 promoter (figure 4.2)

PCR 2: Purified product from step 1 acts as template in a PCR with the hU6 Forward and *NOP10* shRNA 2 primers. 5' to 3' the *NOP10* shRNA 1 primer is composed of an *EcoRI* linker, transcription termination sequence (poly$(T)_6$), 19nt of antisense strand

and 14nt of complementary sequence with the 5' end of *NOP10* shRNA 1. The result is attachment of the shRNA antisense strand and the transcriptional terminator sequence completing the shRNA expression cassette (figure 4.2).

The purified expression cassette was tested for functional *NOP10* silencing by transient transfection into HeLa cells alongside untransfected and mock transfected controls with samples harvested after 24 and 48hr. RNA was extracted, first strand cDNA synthesized and tested for quality and DNA contamination by PCR of the endogenous *ABL* gene (data no shown). *NOP10* silencing was assessed by qPCR using TaqMan® probes (details in appendix B), expression calculated relative to the untransfected control for comparison (described in chapter 2.2.9).

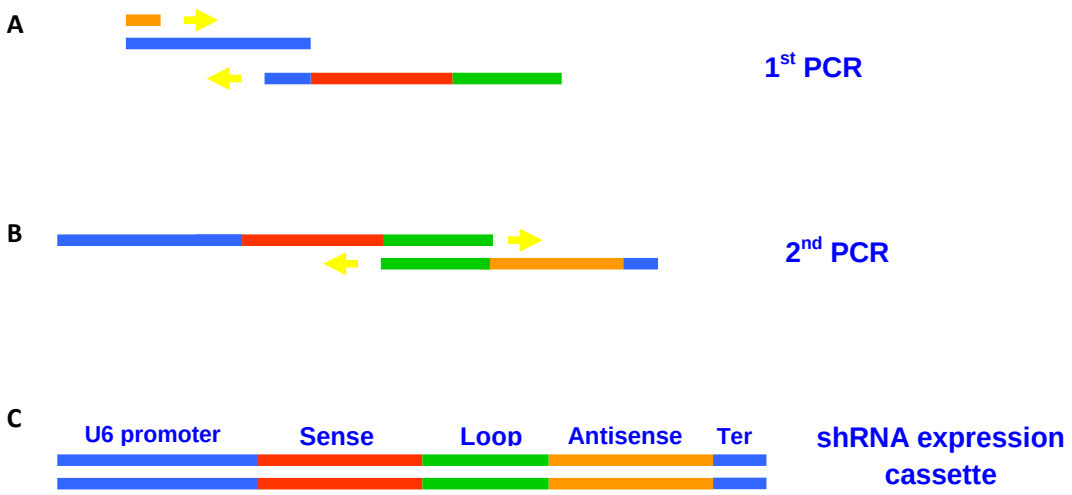


Figure 4.2. PCR synthesis of shRNA expression cassettes. (A) Step 1: PCR links the template U6 promoter and an overlapping reverse primer containing the sense and loop sequence. (B) Step 2: PCR links products from the 1st PCR and a reverse primer containing the loop (for overlap), antisense and terminator (Ter) sequence. (C) The final PCR product consists of the U6 promoter, sense, loop, antisense and termination sequences – the shRNA expression cassette.

4.2.2. Results and Discussion

Following the outlined procedure preparation of the shRNA expression cassette by PCR was successful, products at each stage of the process demonstrated the progressive increase in size from the U6 promoter (346bp) to the finished expression cassette (419bp) (figure 4.3).

Functional testing of the *NOP10* shRNA cassette, however, failed to elicit a decrease in *NOP10* expression. The literature suggests that as little as 100ng of the shRNA cassette is sufficient to generate a silencing effect (Castanotto et al. 2002). Therefore, HeLa cells were transfected with 100, 200, 300 and 400ng of the purified *NOP10* shRNA expression cassette, but after 48hr corresponding *NOP10* expression levels had not been decreased (figure 4.4). In order to verify this result 350ng of the *NOP10* shRNA cassette was transfected a second time and again *NOP10* expression after 24 and 48hr revealed no decreases (data not shown). To investigate this lack of function the expression cassette was sequenced using the hU6 Forward primer and revealed an insertion event in the antisense sequence of the construct (figure 4.5). Repeating 15nt of the upstream sequence this would certainly interfere with stem loop formation of the shRNA and explain the lack of function. The cause of this insertion can only be speculated but its location at the point of overlap between the *NOP10* shRNA 2 primer and the template in step 2 of the process suggests promiscuous annealing or rearrangement.

While this problem could probably be overcome with repeated PCR and optimisation, this finding highlighted PCR errors as a potential problem for this strategy. A single PCR generates a limited amount of the shRNA cassette, often requiring several reactions to be pooled to achieve required amounts. Every new PCR would have to be sequenced for quality control which would be impractical and costly. In light of these problems an alternative strategy was sought.

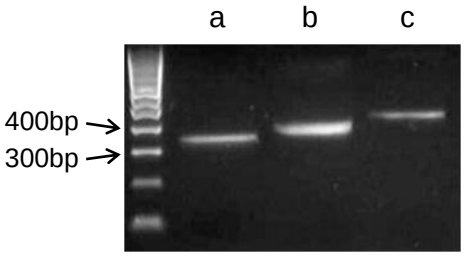


Figure 4.3. PCR formation of the *NOP10* shRNA expression cassette. (a) U6 promoter, (b) Step 1: U6 promoter-sense-loop, (c) Step 2: The finished cassette; U6 promoter-sense-loop-antisense.

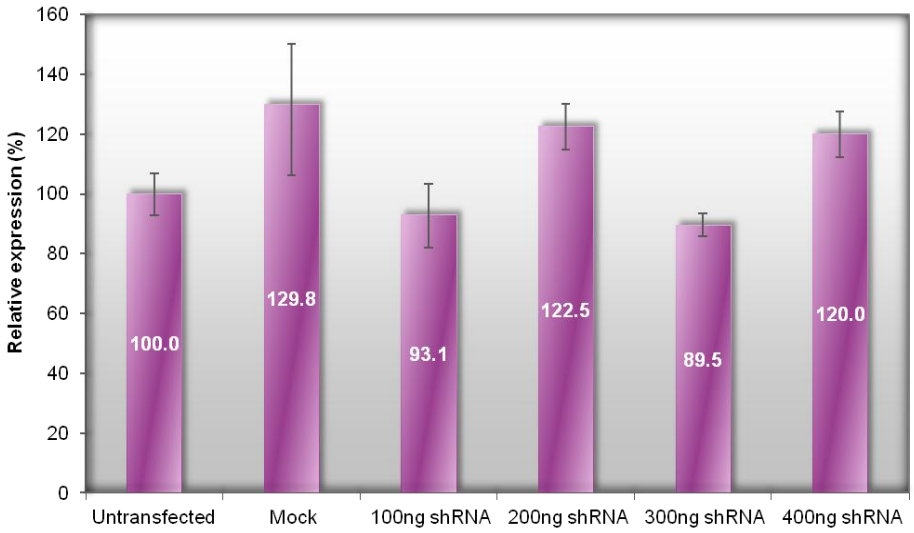


Figure 4.4. *NOP10* expression in HeLa cells transfected with *NOP10* shRNA. Measured 48 hours after transfection with different amounts of the *NOP10* shRNA expression cassette and calculated relative to the untransfected control. Values are the average of triplicate measurements made from a single experiment with error bars marking the SD.

```
          1                                                   50
Expected: ACGCGT-GTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAGCGTG
  Actual: ACGCGTAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAGCGTG

          51                                                  100
Expected: ACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCC
  Actual: ACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCC

          101                                                 150
Expected: CATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAA
  Actual: CATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAA

          151                                                 200
Expected: TTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGA
  Actual: TTAGAATTAATTTGACTGCAAACACAAAGATATTAGTACAAAATACGTGA

          201                                                 250
Expected: CGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTT
  Actual: CGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTT

          251                                                 300
Expected: TAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCT
  Actual: TAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCT

          301                                                 350
Expected: TGGCTTTATATATCTTGTGGAAAGGACGAAACAGCGGCCGCCTCGAGCCC
  Actual: TGGCTTTATATATCTTGTGGAAAGGACGAAACAGCGGCCCCC---AGCCC

          351                                                 400
Expected: ATAAAGGGAACACATTCCACACCAATGTGTTCC----------------CT
  Actual: ATAAAGGGAACACATTCCACACCAATGTGTTCCACACCAATGTGTTCCCT

          401                                 435
Expected: TTATGGGGTCGACTCTAGAAAAAAGGCGCGAATTC
  Actual: TTATGGGGTCGACTCTAGAAAAAAGGCGCGAATTC
```

Figure 4.5. Sequence data for the NOP10 shRNA expression cassette. The actual sequence is aligned against the expected sequence. The sense strand is marked in red, the antisense in orange and the loop in green. Mismatches are highlighted in yellow with dashed lines marking missing/inserted sequence.

4.3. Cloning Strategy

The previous approach attempted to create shRNA expression cassettes that could be cloned into a suitable plasmid. Therefore it was decided to employ a direct cloning strategy. Here the sense - loop - antisense sequence is encoded in a pair of complementary oligonucleotides which are annealed to create a double stranded synthetic DNA molecule. Incorporating restriction sites at each end enables subsequent cloning into the plasmid of choice (figure 4.6). This straight forward cloning method eliminates the potential of introducing sequence errors like those seen in the PCR approach.

4.3.1. Experimental Procedure

Specifically designed for shRNA expression, p*Silencer*™ 2.1-U6 neo (Ambion) (appendix C) was the plasmid chosen for this procedure and came with a pre-prepared negative control p*Silencer* plasmid. Ambion's matching online design tool (http://www.ambion.com/techlib/misc/psilencer_converter.html) was used to generate suitable oligonucleotide sequences for the *NOP10* shRNA insert: *NOP10* shRNA (+) and *NOP10* shRNA (-) (details in appendix B). Corresponding PAGE purified oligonucleotides were purchased (Sigma-Genosys) and annealed to create a synthetic DNA molecule with *BamHI* (5') and *HindIII* (3') 'sticky ends'. This was subsequently ligated into the pre-linearised p*Silencer*™ 2.1-U6 neo plasmid (figure 4.7). Aliquots of the ligation were used to transform DH5α ™ competent cells (Invitrogen) and 6 colonies selected for small scale plasmid preparation and sequencing with the M13 primer pair (details in appendix B).

Functional tests were carried out via a transient transfection experiment in which HeLa cells were co-transfected with the *NOP10* p*Silencer* and pIRES2-EGFP plasmids (1µg each), also co-transfecting with the negative control p*Silencer* accordingly. GFP expression was visualised under fluorescence microscopy after 24hr as an indicator of transfection efficiency. At 48hr post transfection RNA was extracted, first strand cDNA synthesized and tested for quality and DNA contamination by PCR of *ABL* as previous (primers ABL1 and ABL2; data not shown). *NOP10* silencing was assessed by qPCR using TaqMan® probes (details in appendix B), calculated relative to the untransfected control for comparison (described in chapter 2.2.9).

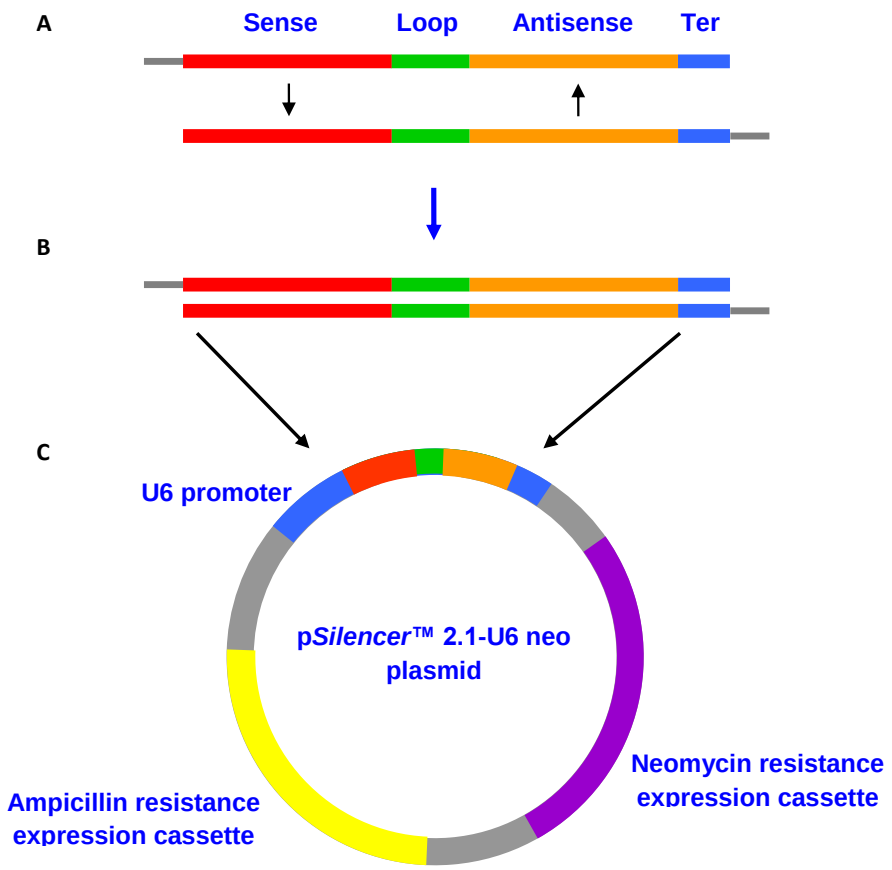


Figure 4.6. A schematic representation of the shRNA cloning process. (A) A pair of complementary oligonucleotides comprising of the sense – loop – antisense – termination (Ter) sequence are annealed. (B) A synthetic DNA molecule encoding the shRNA is created, designed to have 'sticky ends' for directional subcloning. (C) Ligation into p*Silencer*™ 2.1-U6 neo creates the finished shRNA expression plasmid.

To establish stable cell lines expressing the p*Silencer* constructs, HeLa cells were independently transfected with 1µg of the *NOP10* p*Silencer* and negative control p*Silencer* plasmids under neomycin selection (described in chapter 2.2.2.2). At the first passage (~3 weeks post transfection) samples were used to extract RNA and first strand cDNA synthesized to evaluate *NOP10* silencing by qPCR as described earlier. Genomic DNA was also extracted and PCR used to confirm integration of the p*Silencer* construct (p*Silencer* FWD and REV primers, details in appendix B). Simultaneously, limiting dilutions were used to create clonal cell lines for both and *NOP10* silencing was verified once established (as previous).

4.3.2. Results

Oligonucleotide annealing and cloning into the p*Silencer*™ 2.1-U6 neo plasmid was straight forward and highly successful with all of the 6 clones sequenced found to contain the correct shRNA insert (figure 4.7).

```
     1                                                                   70
Exp: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C1: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C2: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C3: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C4: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C5: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
 C6: GGATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAAAGCTT
```

Figure 4.7. Sequence data for *NOP10* p*Silencer*. The shRNA insert of clones 1-6 (C1-6) are aligned against the expected sequence (Exp), sense strand sequence is marked in red, antisense in orange and the loop in green.

Functional tests in HeLa cells transfected with the *NOP10* p*Silencer* plasmid demonstrated effective *NOP10* silencing. Visual GFP analysis estimated transfection efficiency to be between 50-60% after 24hr (figure 4.8). Corresponding to this level of transfection, *NOP10* levels measured at 48hr had decreased to 49.2% under the *NOP10* p*Silencer* while remaining at a higher 143.9% under the negative control p*Silencer* (figure 4.9).

Transfection and neomycin selection was subsequently used to generate stable polyclonal cell lines expressing these constructs. At 3 weeks post transfection genomic integration of the p*Silencer* constructs was verified by PCR (figure 4.10) and analysis of *NOP10* expression confirmed suppression in the *NOP10* p*Silencer* cell line, levels decreased to 19.6% compared to an untreated control (figure 4.11). However, *NOP10* expression was also reduced to 48.3% in the negative control p*Silencer* cell line (figure 4.11), an effect that had not been seen in the earlier transient test of this construct. To exclude the possibility of a mistake the integrated p*Silencer* constructs were PCR amplified a second time and sequenced using the same primers, (p*Silencer* FWD and REV). Results confirmed that each line contained the correct construct (data not shown).

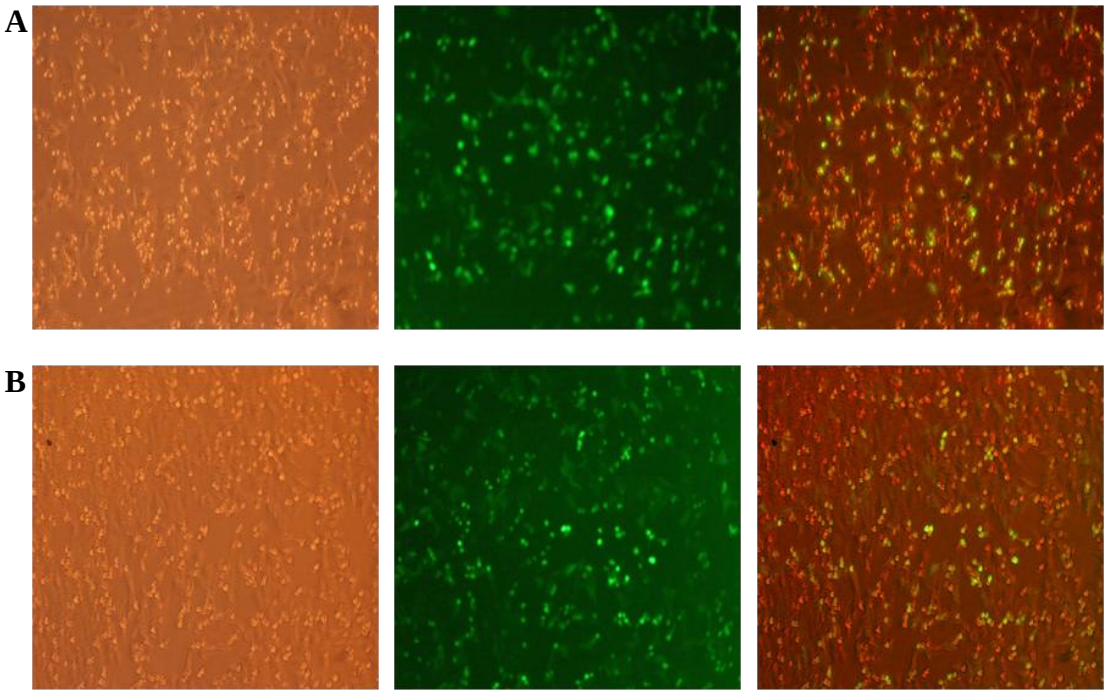


Figure 4.8. GFP expression in HeLa cells 24 hours post transfection. Corresponding bright field (left), fluorescence (middle) and overlaid (right) images of (A) *NOP10* p*Silencer* transfected cells. (B) Negative control p*Silencer* transfected cells.

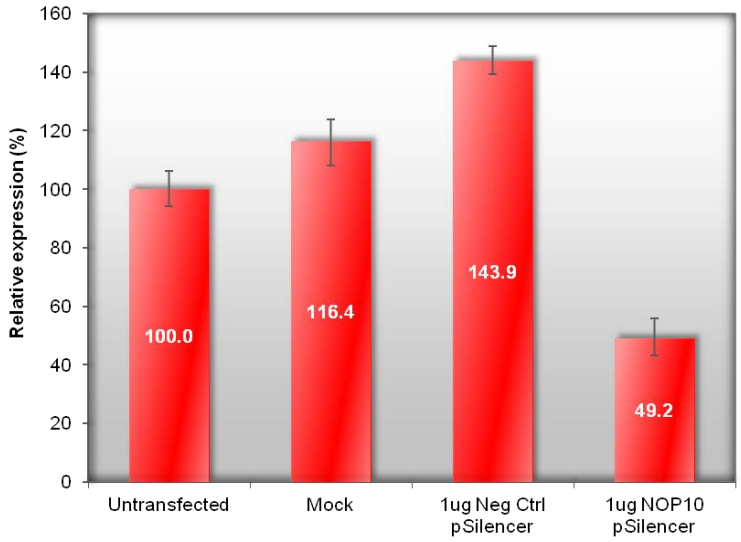


Figure 4.9. *NOP10* expression in HeLa cells transfected with p*Silencer* plasmids. Measured 48 hours post transfection with 1µg of the *NOP10* p*Silencer* and negative control p*Silencer* plasmids, expression was calculated relative to the untransfected control. Values are from a single experiment, error bars marking the SD between assay triplicates.

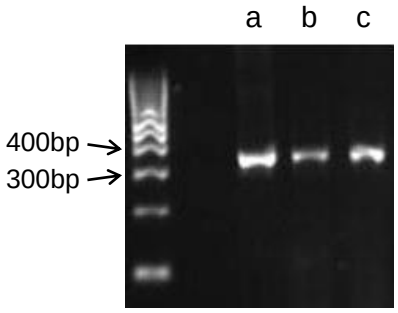


Figure 4.10. PCR confirming genomic integration of the p*Silencer* constructs. (a) *NOP10* p*Silencer* plasmid (positive control; 330bp). (b) negative control p*Silencer* cell line. (c) *NOP10* p*Silencer* cell line.

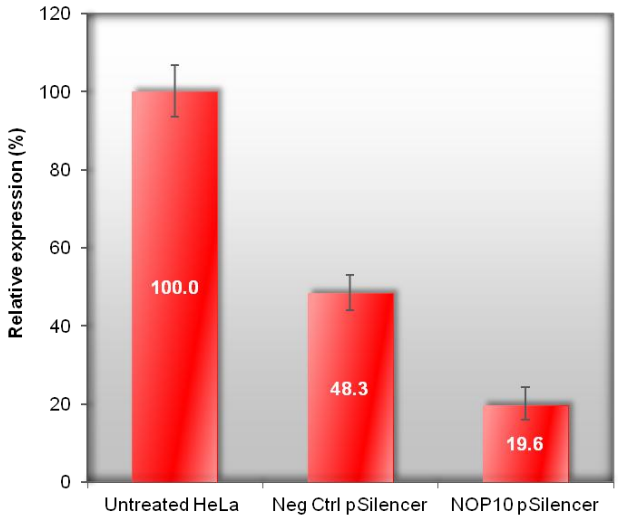


Figure 4.11. *NOP10* expression in the p*Silencer* cell lines. Measured in the negative control p*Silencer* and *NOP10* p*Silencer* cell lines 3 weeks post transfection and calculated relative to an untreated HeLa control cell line. Values are from single cell lines, error bars marking the SD between assay triplicates.

After a couple of attempts to generate clonal cell lines by limiting dilution, five *NOP10* p*Silencer* and three negative control p*Silencer* cell lines were eventually produced. To begin functional analysis *NOP10* expression and TERC levels were evaluated in each (9-10 weeks post transfection at this stage) (figure 4.12). Compared to an untreated HeLa control, *NOP10* transcript levels were substantially suppressed in all five of the *NOP10* p*Silencer* clones (average expression below 30%). However, corresponding with the polyclonal parent line, *NOP10* levels were also reduced in the negative control p*Silencer* clones. Corresponding TERC levels were reduced in all cases, though subtle differences were observed between the *NOP10* p*Silencer* clones (16% or less) and the negative p*Silencer* clones (20-30%).

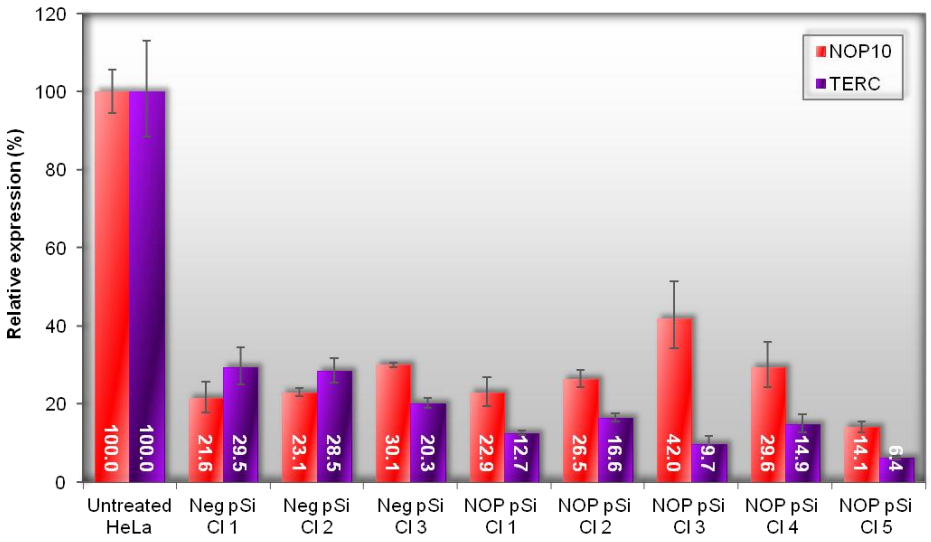


Figure 4.12. *NOP10* and TERC levels in clonal p*Silencer* cell lines. Measured in negative control p*Silencer* (Neg pSi) and *NOP10* p*Silencer* (NOP pSi) cell lines and calculated relative to the untreated HeLa control. Values are from single cell lines, error bars marking the SD between assay triplicates.

4.3.3. Discussion

After changing the strategy to a direct cloning approach, the *NOP10* p*Silencer* plasmid was created and confirmed to be functional. Given that the sense and antisense sequence of the shRNA matches that of *NOP10* siRNA 3, used in the transient experiments (chapter 3), it is reasonable to assume that both would be equally effective. Yet, initial transient tests of the plasmid elicited only a 50% decrease in *NOP10* expression after 48hr, unimpressive in contrast to the 90%+ decrease after only 24hr using the siRNA. There are important factors to be considered in this comparison though. Firstly, siRNA takes effect much faster than shRNA. RNAi using shRNA requires DNA expression and subsequent downstream processing to produce a functional siRNA (outlined in figure 4.1). These steps are not involved when using siRNA directly and allow it to take effect more swiftly. Secondly, differences in transfection efficiency. Only cells containing the siRNA or shRNA can be subjected to their suppressive effects; therefore, silencing efficiency can never exceed the transfection efficiency. Vastly different in size, the smaller siRNA (19bp) will transfect with far greater efficiency than the larger p*Silencer* plasmid (4585bp). Indeed, the observed 90%+ decrease in *NOP10* expression when using the siRNA would indicate that at least 90% of the cell population was transfected. In comparison, visual GFP analysis (estimated by eye) suggests *NOP10* p*Silencer* transfection efficiency to be no more than 50-60% which correlates with the level of *NOP10* silencing observed. Moreover, this correlation suggests that *NOP10* p*Silencer* is highly effective with delivery into the cell being the limiting factor. On this basis it was expected that as the number of cells positive for the *NOP10* p*Silencer* increased, *NOP10* expression levels would decrease further.

This was confirmed in the stable cell line generated under neomycin selection where *NOP10* expression fell to less than 20% in the *NOP10* p*Silencer* cell line. Unfortunately this was marred by *NOP10* levels being reduced to less than 50% in the negative control p*Silencer* cell line as well, though this had not been observed in the initial transient test of this construct. In view of this contradiction, both cell lines were confirmed to have the correct p*Silencer* construct integrated in their genome by PCR and sequencing. Exacerbating the issue, clonal cell lines derived from the negative control p*Silencer* cell line exhibited further reductions in *NOP10* expression, comparing closely to those derived from the *NOP10* p*Silencer* cell line (between 20-

30%). Corresponding to this, TERC levels were also reduced in all the lines. At this stage it was clear that the effects observed in the *NOP10* p*Silencer* lines could not be attributed to *NOP10* p*Silencer* specifically and consequently the experiment was terminated.

On reflection this strategy had not proved to be straight forward and required more than one attempt, particularly the limiting dilutions. The time taken to conduct these experiments and the degree of uncertainty involved was a concern, especially with regards to the selection process and clonal expansion. This is effectively an indeterminable 'black box' period before the experiment can truly begin and is of variable duration depending on the recovery and proliferation of the cells, on this attempt 9-10 weeks. Having already spent significant time on this strategy careful evaluation was warranted before deciding how to proceed.

Indications were that the negative control p*Silencer* exerted a detrimental effect on *NOP10* and TERC levels that was not observed transiently but manifested gradually over time. This was quite different to the targeted effect of the *NOP10* p*Silencer* that was evident from 48hr suggesting a subtle and cumulative indirect effect. Off targeting in RNAi experiments has been discussed in chapter 3.1 and could potentially be the issue here. Another possibility is that this is the result of something more fundamental to the experimental process, perhaps the neomycin selection. Unfortunately the experiment lacks a suitable neomycin resistance only (or mock) control to distinguish the two possibilities. However, this method of generating stable cell lines is not uncommon and the use of antibiotic selection unavoidable. Instead, investigating a different negative control p*Silencer* presented a more feasible option but, with no substitute available, this would have to be designed and produced. Alternatives were available in other commercial shRNA expression systems, the expanding market of RNAi resources also offering pre-designed and prepared shRNA expression vectors against an increasing number of targets. Switching to an alternative system presented a quick and simple solution to replacing the negative control without the complication of further cloning. Importantly, the *NOP10* p*Silencer* could also be substituted with the same ease, as would be necessary for maintaining consistency in the experiment. Therefore, this strategy was reconsidered with a view to changing direction and implementing a different shRNA expression system.

4.4. Lentiviral Strategy

Lentiviral infection provides a highly efficient method for delivery and stable integration of expression constructs into cells that avoids the need for transfection, a potentially harsh process not suitable for all cell types. On this basis lentiviral vectors are favoured commercially for pre-designed and prepared shRNA expression constructs and were consequently chosen to replace the p*Silencer* plasmids and the negative control used previously.

The initial aim had been to carry out long term knockdown experiments for dyskerin, NOP10, NHP2 and GAR1. However, efforts so far had been problematic and time consuming and so at this point it was decided to focus on the long term knockdown of just a one of these targets. Of the four, dyskerin is the most commonly mutated in DC and posed the more significant target. Therefore, efforts were refocused on the stable suppression of *DKC1* instead.

4.4.1. Experimental Procedure

The pre-designed and validated lentiviral expression vectors MISSION® *DKC1* shRNA and MISSION® Non-Target shRNA (Sigma) (appendix C) were chosen for this experiment. The pre-prepared plasmids were transformed into DH5α™ competent cells (Invitrogen) and used for large scale plasmid preparations followed by transfection into HEK293T cells for lentiviral packaging (described in chapter 2.2.3). The lentivirus was then concentrated from the culture medium to produce aliquots of each virus (described in chapter 2.2.4) which were independently used to transduce HeLa cells and generate stable cell lines under puromycin selection (described in chapter 2.2.5). Cultured for 18 weeks, cell counts were used to monitor population doublings and samples were taken at regular intervals for the analysis of *DKC1* expression, TERC accumulation, TRAP activity and H/ACA RNA accumulation, as described for the transient experiments in chapter 3. To investigate for impacts on telomere maintenance, genomic DNA samples were used to assess telomere lengths in each cell line by southern blot (described in chapter 2.2.19). Measured against two pre-determined reference samples that were used as size standards, the values obtained were subsequently adjusted to account for subtelomeric DNA (~7.8kb) also encompassed within the fragment.

4.4.2. Results

Preparation of the lentivirus proved to be straight forward and was successfully used to create polyclonal *DKC1* shRNA and non-target control (NTC) shRNA stable cell lines. Comparatively, the *DKC1* shRNA cell line took longer to recover from the selection process and reach first passage, initially suffering a high level of cell death in a crisis period that may indicate poor viral transduction efficiency. As a result *DKC1* expression levels could not be verified until 5 weeks post transduction, at which point analysis revealed a substantial decrease to 17.3% in the *DKC1* shRNA cell line compared to the untreated control (figure 4.13). Importantly, although decreased, *DKC1* expression in the NTC shRNA cell line remained much higher at 66.1% (figure 4.13). Having confirmed effective *DKC1* suppression the cultures were maintained for a further 13 weeks (18 weeks in total) during which time population doublings were monitored and samples taken for analysis at regular intervals. Over this 13 week period the growth of the *DKC1* shRNA cell line was comparatively slow and impaired, undergoing fewer population doublings than both the NTC shRNA and untreated HeLa cell lines which grew similarly to each other (figure 4.14).

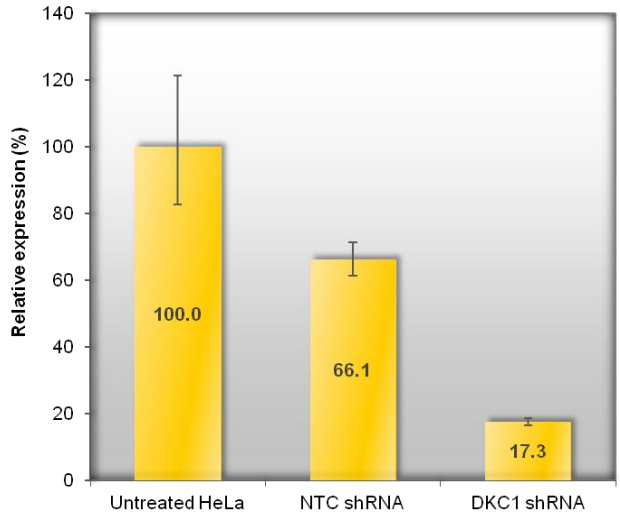


Figure 4.13. *DKC1* expression levels. Measured 5 weeks post transduction, expression in the NTC shRNA and *DKC1* shRNA cell lines was calculated relative to an untreated HeLa control. Values are from single cell lines, error bars marking the SD between assay duplicates.

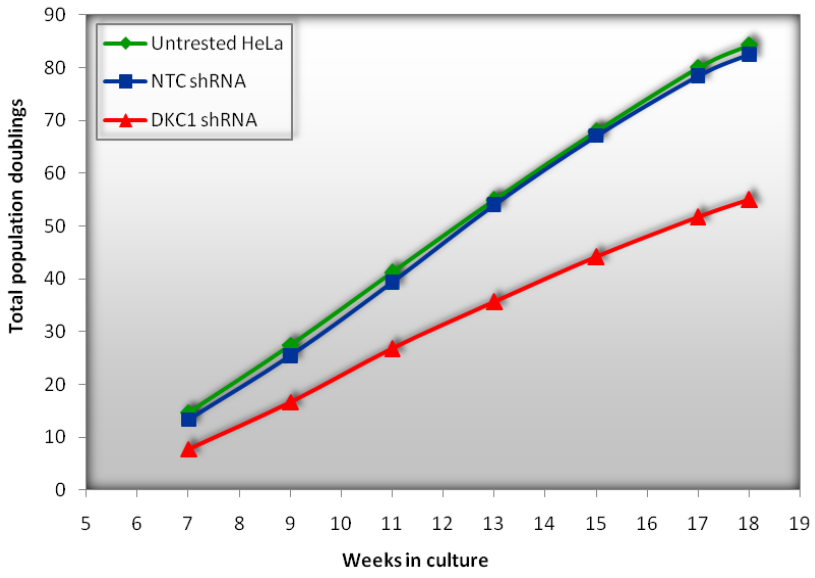


Figure 4.14. Growth curves for the shRNA cell lines. Proliferation over the 13 week period from weeks 5 to 18 post transduction is compared for the untreated HeLa, NTC shRNA and *DKC1* shRNA cell lines.

4.4.2.1. *DKC1* Suppression

Measured at 10, 13 and 16 weeks post transduction, *DKC1* expression was consistently reduced to below 30% in the *DKC1* shRNA cell line compared to the untreated control (figure 4.15). Corroborating with the analysis at 5 weeks, this confirmed a sustained shRNA effect suppressing *DKC1* expression, and by implication dyskerin, in this cell line. It is noteworthy that in this line there was suggestion of a slow, steady increase in *DKC1* expression from 18% at 10 weeks to 28% at 16 weeks. This could be simple fluctuation but considering the polyclonal nature of this cell line it may also be evidence of selective pressure against shRNA expression, the course of the culture favouring the cells least affected i.e. less potent suppression. Certainly, the slow growth of this cell line strongly suggests that the long term suppression of *DKC1* conveys a disadvantage that would explain any selective disadvantage.

Conversely the NTC shRNA cell line exhibits a gradual decline in *DKC1* expression from 72% to 45% over the term of the culture (figure 4.15). Though this echoes prior issues with the negative p*Silencer* experiments the impact here is somewhat less dramatic. The striking decrease observed in the untreated control at 13 weeks is inconsistent and most likely an erroneous result.

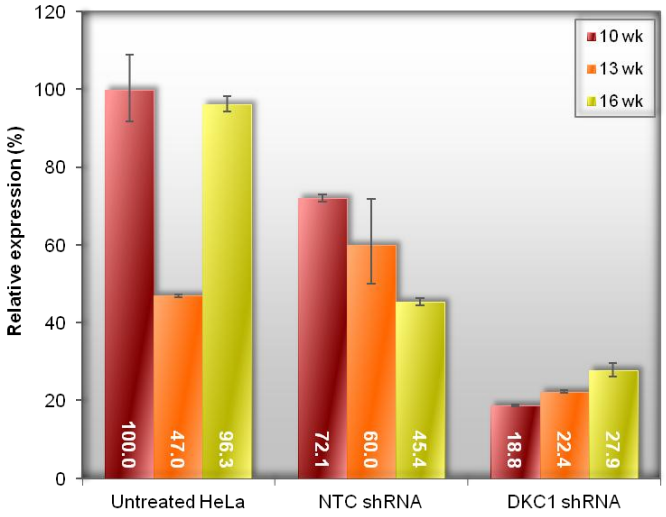


Figure 4.15. *DKC1* expression in the shRNA cell lines. Measured 10, 13 and 16 weeks post transduction, expression in the NTC shRNA and *DKC1* shRNA cell lines was calculated relative to the untreated HeLa control at 10 weeks. Values are from single cell lines, error bars marking the SD between assay duplicates.

4.4.2.2. TERC Levels

Supporting the findings of the transient knockdown experiment and consistent with a decrease in dyskerin, TERC levels were substantially reduced in the *DKC1* shRNA cell line at each time point (figure 4.16a). Strikingly, TERC levels were much higher at 10 weeks (49.6%) than at 13 and 16 weeks (12 and 20.5% respectively) despite comparably low *DKC1* expression at all three time points. To investigate this further TERC levels were also measured at the earlier 5 week time point, the results showing a decrease to less than 25% (figure 4.16b) that is more consistent with the analysis at 13 and 16 weeks. Therefore, it appears that the TERC level at 10 weeks is somewhat irregular may be erroneous.

In contrast, TERC levels in the NTC shRNA cell line were consistently higher showing a slow steady increase from 71% at10 weeks to 87% at 16 weeks (figure 4.16a). Measured at 5 weeks TERC levels in this cell line were 81% (figure 4.16b) suggesting no significance to the observed increase other than random fluctuation.

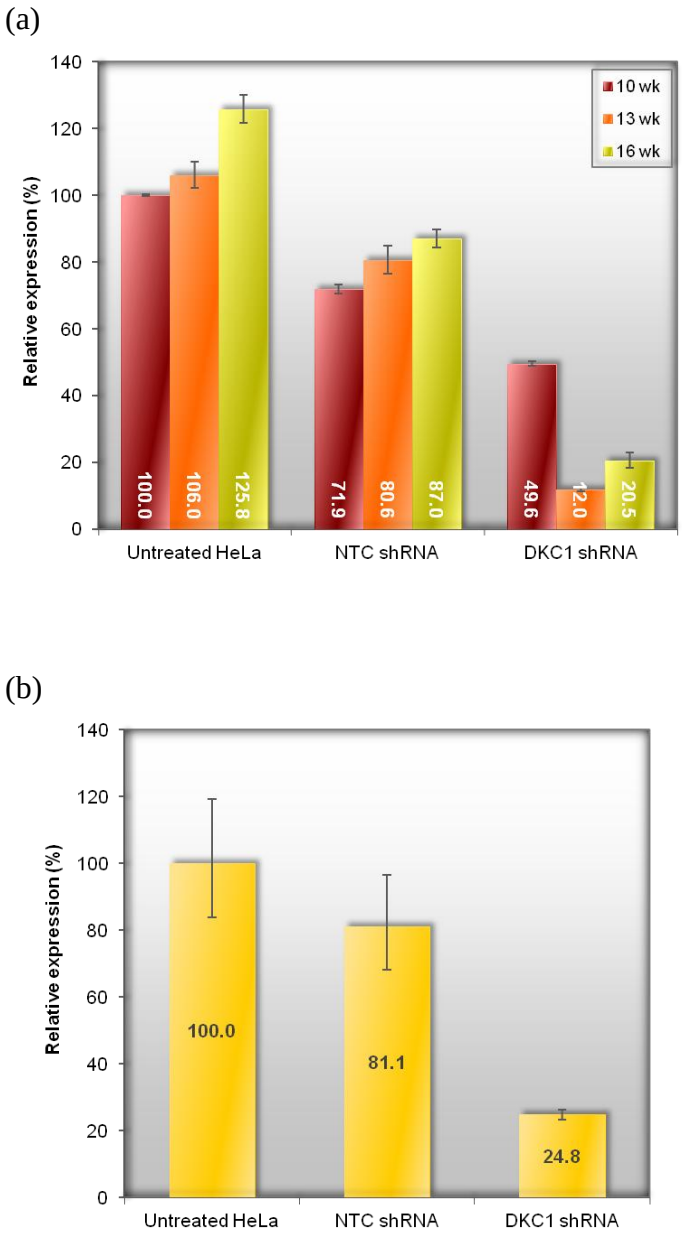


Figure 4.16. TERC levels in the shRNA cell lines. (a) Measured 10, 13 and 16 weeks post transduction, levels in the NTC shRNA and *DKC1* shRNA cell lines were calculated relative to the untreated HeLa control at 10 weeks. (b) Similarly measured at 5 weeks post transduction. Values are from single cell lines, error bars marking the SD between assay duplicates.

4.4.2.3. TRAP Activity

Consistent with the reduced TERC levels, TRAP analysis at each time point demonstrates that activity was severely and persistently diminished to below 30% in the *DKC1* shRNA cell line compared to the untreated control (figure 4.17). In comparison TRAP activity in the NTC shRNA cell line was remarkably higher at 103-179% and although displaying a downward trend over the time course, remained within a comparable range to that of the untreated HeLa control (figure 4.17).

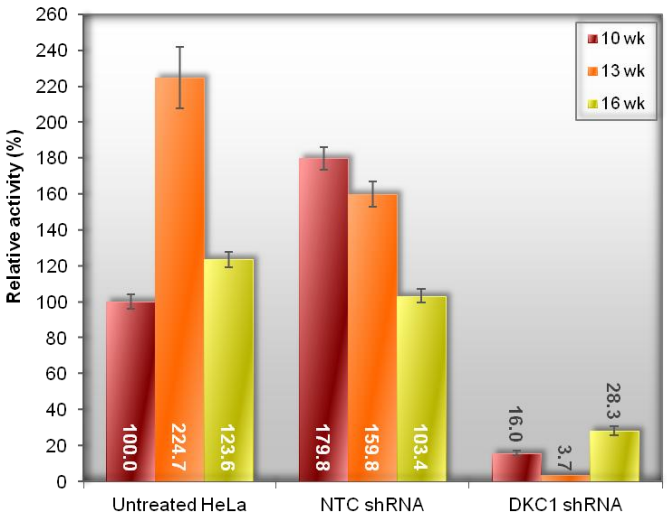


Figure 4.17. TRAP activity in the shRNA cell lines. Measured 10, 13 and 16 weeks post transduction, activity in the NTC shRNA and *DKC1* shRNA cell lines was calculated relative to the untreated HeLa control at 10 weeks. Values are from single cell lines, error bars marking the SD between assay triplicates.

4.4.2.4. Telomere Length

Southern blot analysis of telomere length at 7, 11 and 18 weeks post transduction failed to yield a result for week 7 (attributed to sample quality and loading issues) but identified measurable bands for weeks 11 and 18 (figure 4.18). It should be noted that, seeded from the same stock culture, all cell lines had equivalent telomere lengths when the experiment began.

Comparing telomere length across the three lines after 18 weeks shows that over the course of the experiment telomeres shortened comparably in both the NTC shRNA and *DKC1* shRNA cell lines relative to the untreated HeLa control (4, 3.5, 5.7kb, respectively). However, making this comparison at 11 weeks only indicates short telomeres in the *DKC1* shRNA cell line, telomeres in the NTC shRNA cell line maintained at a similar length to the untreated HeLa control at this time point (3.8, 4.8, 5.2kb, respectively). Therefore it is evident that the onset and dynamics of telomere shortening were different in each of the lines, commencing sooner in the *DKC1* shRNA cell line. What's more, this cell line exhibited comparable telomere lengths at both 11 and 18 weeks indicating that they were maintained, albeit shorter, over this period and suggesting an early transient period of telomere shortening that eventually stabilised.

In contrast, the untreated HeLa cell line showed no sign of telomere shortening during the course of the experiment.

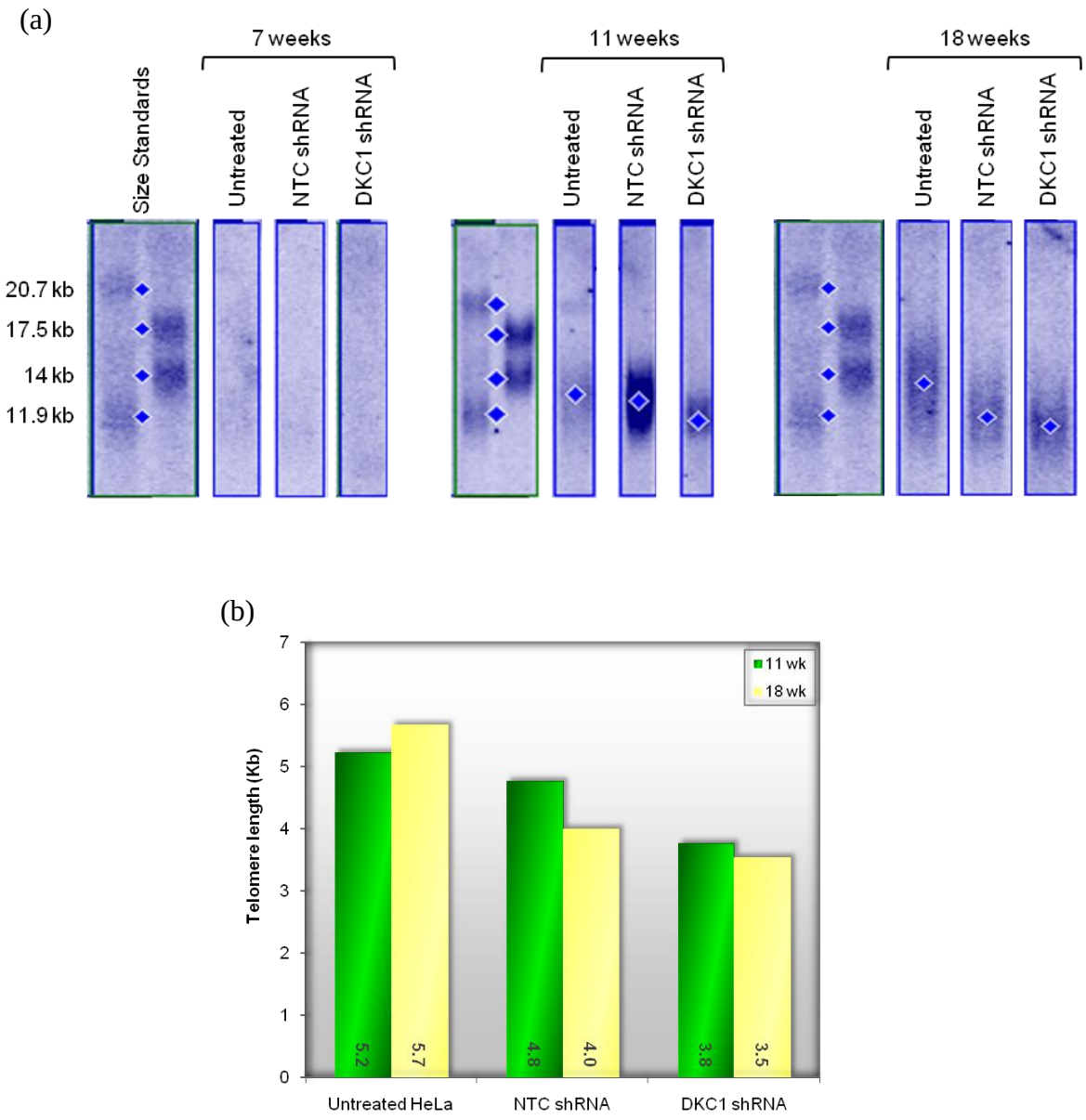


Figure 4.18. Telomere length in the shRNA cell lines. (a) Southern blot analysis of the untreated HeLa control, NTC shRNA and *DKC1* shRNA cell lines at 7, 11 and 18 weeks post transduction. Diamonds mark the reference points used for measurement against the size standards. **(**b) Adjusted telomere lengths at 11 and 18 weeks post transduction.

4.4.2.5. H/ACA RNA Levels

Issues pertaining to sample quality resulted in analysis of the H/ACA RNAs: TERC, U19, U64, U68, U70, and the C/D box RNAs U97 and U3, at 10 and 16 weeks only (figure 4.19).

Proving to be a more reliable data set than that obtained for the transient experiments (chapter 3), the results here reaffirm that, at both time points, TERC levels were substantially reduced in the *DKC1* shRNA cell line compared to the untreated HeLa and NTC shRNA cell lines. Similarly, at 10 weeks the *DKC1* shRNA cell line also exhibited substantial decreases in levels of the other H/ACA RNAs that were not seen in the NTC shRNA cell line, most strikingly U68 and U70. Behaving differently to TERC, however, these decreases did not persist at 16 weeks, by which time only U68 was still reduced but less substantially. Subtle decreases were still evident in U19 and U70 as well, but similar reductions present in the NTC shRNA cell line render this less convincing. In contrast, levels of the C/D box RNA U97 were maintained at both time points and appear to be unaffected.

It is important to note that in the NTC shRNA cell line the H/ACA RNAs were largely unaffected. However, as previous, levels of the U3 C/D box RNA proved to be erratic and were discounted.

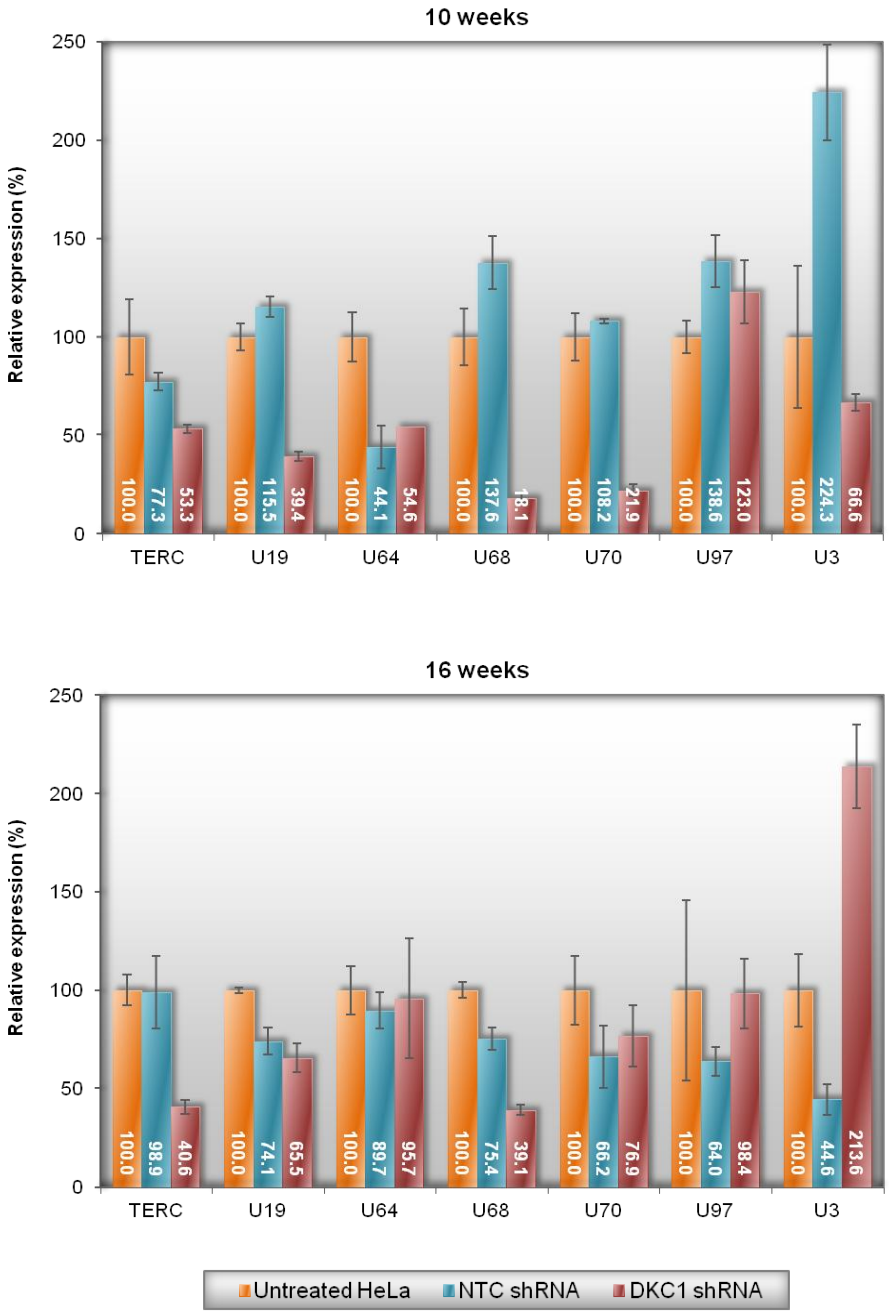


Figure 4.19. H/ACA RNA levels in the shRNA cell lines. Measured at 10 and 16 weeks post transduction, levels in the NTC shRNA and *DKC1* shRNA cell lines were calculated relative to those in the untreated HeLa control. Values are from single cell lines, error bars marking the SD between assay duplicates.

4.4.3. Discussion

After encountering serious issues attempting a stable knockdown experiment using the p*Silencer* system, the changes implemented here proved effective and a system for investigating the stable knockdown of dyskerin was successfully achieved. Lentiviral MISSION shRNAs were used to create the *DKC1* shRNA and NTC shRNA stable cell lines and their *DKC1* expression levels were assessed. The pre-validated shRNA in the *DKC1* shRNA cell line proved to be highly effective and suppressed *DKC1* expression to below 30% for the duration of the experiment, implicating a corresponding reduction in dyskerin levels. Importantly, although the NTC shRNA did display a gradual decline to 45% at week 16, for most of the experiment (13+ weeks) levels were maintained sufficiently higher at over 60%.

Consistent with the findings of the transient knockdown in chapter 3 and similar transient studies elsewhere (Montanaro et al. 2006; Montanaro et al. 2008), subsequent analysis confirmed a sustained and substantial negative impact on TERC levels and TRAP activity corresponding to the knockdown of dyskerin in the *DKC1* shRNA cell line. Reaffirming the critical role of dyskerin, it is evident that telomerase stability and function were compromised in this cell line. As core machinery for telomere maintenance, over time this would be expected to have severe ramifications on telomere length. Investigating this, telomeres were found to be comparably short in both the *DKC1* shRNA and NTC shRNA cell lines at the end of the experiment (18 weeks). Nonetheless, it appears that telomeres shortened much earlier in the *DKC1* shRNA cell line (prior to 11 weeks) than in the NTC shRNA cell line (post 11 weeks). Therefore, telomeres were affected differently in the *DKC1* shRNA cell line compared to the controls which supports the hypothesis of a comparative defect in their maintenance. However, with telomeres maintained (albeit shorter) between 11 and 18 weeks it would appear that this manifested only transiently and then stabilised. Reconciling this with the sustained telomerase impairment in this cell line is difficult: why would this elicit only a transient rather than equally sustained downstream impact? One explanation would be that an alternative mechanism of telomere maintenance was activated in response to their shortening.

Cells that proliferate indefinitely, such as immortalised cell lines, will predominantly maintain their telomeres using telomerase. Nevertheless, there are telomerase negative

cell lines that achieve this via the alternative lengthening of telomeres (ALT) pathway, a process of DNA recombination that is not fully understood (Henson et al. 2002). The yeast model in which this was first characterised is not dissimilar to the *DKC1* shRNA cell line. Mutated to disrupt telomerase activity, telomere shortening and replicative senescence ensued causing the yeast culture to go into crisis. From this emerged two types of rare survivors that could be differentiated by growth and telomere length. Type I survivors grew slowly and stably maintained a short homogeneous population of telomeres. Type II survivors grew normally and exhibited a dynamically maintained heterogeneous telomere population that ranged from exceptionally long to very short and were prone to abrupt changes as well as the gradual shortening that marked the underlying telomerase defect. Subsequent work demonstrated that each survivor type was dependent on a different recombination pathway, type I mediated through *RAD51* and type II through *RAD50*. These pathways are not fully defined and what decides a type I or II survivor outcome is unknown (reviewed in Henson et al. 2002; Lundblad 2002; Nabetani and Ishikawa 2011).

Similarly afflicted with impaired telomerase activity, parallels to this yeast model can be identified in the *DKC1* shRNA cell line. Compounded and masked by the antibiotic selection process, the excessive level of cell death encountered soon after viral transduction marks a crisis period in this cell line with the prolonged recovery highlighting how few cells survived. Emerging from this, the *DKC1* shRNA cell line grew comparatively slower than the control cell lines and maintained stably short telomeres for the remainder of the experiment. It is therefore plausible that the effects of the dyskerin knockdown did cause telomere shortening as suggested, reasoning that telomere length rapidly became critical and stimulated levels of replicative senescence and cell death that sent the culture into crisis very early on. It is interesting to futher speculate that perhaps the small number of cells that survived did so by activating ALT pathways to circumvent the telomerase deficiency and restore telomere maintenance, slowly reconstituting the culture and the *DKC1* shRNA cell line to resemble a type I survivor from the yeast model.

Human cell lines that utilise the ALT pathway (commonly referred to as ALT cells) typically resemble type II survivors and it was speculated that unlike yeast they

possessed only this pathway (Henson et al. 2002; Lundblad 2002; Nabetani and Ishikawa 2011). However, ALT cells resembling the type I survivor type have been observed. (Fasching et al. 2005; Marciniak et al. 2005). In the yeast model, the slow growth of type I survivors gave type II survivors a selective advantage so that type I survivors were eventually lost from the culture (Lundblad 2002). Perhaps, both pathways exist in human ALT cells with this selection explaining the predominance of the type II survivor phenotype. In this scenario, for a type I survivor phenotype to dominate the type II pathway must somehow be repressed or blocked. If the ALT pathway is active in the *DKC1* shRNA cell line then its resemblance to a type I survivor may suggest some importance for dyskerin in the type II survivor pathway. Interestingly, studies have suggested that dyskerin is involved in the DNA damage response pathway with some mutations increasing localisation of DNA damage foci to the telomeres (Gu et al. 2008; Kirwan et al. 2011), perhaps there is a link here.

ALT cells are naturally telomerase negative, unlike the *DKC1* shRNA cell line in which telomerase is simply depleted. There is uncertainty regarding what influence telomerase activity has on ALT pathways, some studies suggest a possible inhibitory affect that is discounted in other studies (reviewed Lundblad 2002). What's more, these studies all involve the ectopic expression of telomerase in ALT cells, different again from the *DKC1* shRNA cell line. However, one study has investigated the ALT pathway in telomerase deficient mice (through deletion of the telomerase RNA gene) and report evidence of only the type I survivor phenotype. Interestingly, this study also reported evidence of increased recombination at the telomeres of a DC patient with a *TERT* mutation and very short telomeres. Although inconclusive, this suggests that the ALT pathway may be activated in DC, showing resemblance to the type I survivor phenotype (Morrish and Greider 2009). The common element in the mouse study and the *DKC1* shRNA cell line is that telomerase is expressed, albeit functionally disrupted, in these cell types. If ALT pathways are active here, perhaps this suggests that the expression of telomerase inhibits the type II survivor pathway in mammalian cells. More research is needed before these questions can be answered but it is interesting to consider a role for the ALT pathway in DC.

The sustained suppression of dyskerin was also shown to have a variable negative impact on other H/ACA RNA levels, further supporting the results of the transient

knockdown experiment and studies elsewhere (Lafontaine et al. 1998; Dez et al. 2001; Mochizuki et al. 2004) Surprisingly, while the impact on TERC persisted this was not the case for the other H/ACA RNAs measured, evident at 10 but not 16 weeks (with the exception of U68). The reason for this is unclear and may reflect compensatory mechanisms such as increased expression levels. The literature offers very little for comparison on this, most dyskerin knockdown studies have been transient and not all of them investigated H/ACA RNA levels. Therefore, while other studies can corroborate the variable impact initially seen in this study, multiple analyses over a prolonged period were not carried out (Lafontaine et al. 1998; Dez et al. 2001; Mochizuki et al. 2004; Alawi et al. 2011). Functionally, H/ACA RNAs act as guides for pseudouridylation which is catalysed by dyskerin. Unfortunately it was not possible to assay this process in these experiments but transient studies elsewhere have shown that depleting dyskerin impairs global pseudouridylation and rRNA processing (Lafontaine et al. 1998; Montanaro et al. 2006). Therefore, it would be expected that the sustained suppression of dyskerin in these experiments would have a similar impact. The affect on protein biosynthesis and cell survival/proliferation is uncertain. In yeast the depletion of dyskerin has no bearing on telomere maintenance but is associated with a decrease in cell proliferation and growth arrest (Lafontaine et al. 1998). Similarly, the conditional knock out of *DKC1* in adult mice has been shown to severely disrupt ribosome biogenesis and cell proliferation before telomere shortening occurs (Ge et al. 2010). However, studies on human cell lines are inconsistent in their findings. Not dissimilar to the *DKC1* shRNA cell line, one study using telomerase positive prostate cancer cells reports a rapid cessation in proliferation and growth from which only a few cells survived after 10 days but neither H/ACA RNA levels, rRNA processing or telomerase function were investigated (Sieron et al. 2009). Another study using telomerase negative ostesarcoma cells reported only a mild decrease in proliferation and cell numbers after 6 days, also showing a sustained decrease in H/ACA RNA levels but only a transient and subtle disruption to rRNA processing (Alawi et al. 2011). In contrast, a study using a breast cancer cell line reported protein biosynthesis to be unaffected after 4 days of dyskerin knockdown but did not comment on changes in cell number or proliferation (Montanaro et al. 2010). Therefore, although it is expected that pseudouridylation activity would have been affected in the *DKC1* shRNA cell line, it

is difficult to ascertain the extent of the impact and its contribution to the growth impairment observed in this cell line.

In conclusion, these experiments have reiterated dyskerin's importance for the integrity and function of telomerase and demonstrated that in this capacity it is essential for maintaining telomere length over time. Nevertheless, it is also evident that some cells are able to circumvent the telomerase defect and maintain short telomeres, possibly by initiating ALT pathways. The proliferation of these cells is impaired, however, and their long term viability remains to be determined. It is inevitable that pseudouridylation activity will have also been impaired in these cells, but unfortunately it was not possible to determine the extent of the impact and any bearing on cell growth and proliferation remains to be verified.

Chapter 5. *In vitro* analysis of DC mutations.

5.1. NOP10 and NHP2 Mutations

5.1.1. Introduction

Discussed in chapter 1, DC patients have been found to harbour mutations affecting core components of the telomerase complex, predominantly dyskerin, TERC and to a lesser extent TERT. Unpublished prior to commencing this study, genetic analysis conducted by Dr Amanda Walne had also identified mutations affecting two other telomerase components, NOP10 and NHP2, in previously uncharacterised DC patients.

A homozygous *NOP10* missesnse mutation, c.100C>T (p.R34W), was identified in a large consanguineous family (DCR148) comprising three affected siblings who presented with reticular skin pigmentation, nail dystrophy, and hypocellular bone marrow. Both parents and three other siblings were heterozygous for this base change and remained asymptomatic, demonstrating that the homozygous mutation segregated with the disease (figure 5.1a) (Published Walne et al. 2007). Three different *NHP2* missense mutations were identified in two unrelated patients. A homozygous c.415T>C (p.Y139H) mutation was identified in a young male who presented with nail dystrophy, thrombocytopenia, testicular atrophy, growth and mental retardation, liver cirrhosis and also suffered opportunistic infections (family DCR096). Both parents were heterozygous for this mutation and asymptomatic (figure 5.1b). Compound heterozygous mutations, c.376G>A (p.V126M) and c.460T>A (p.X154RextX*52) were identified in another young male who presented with nail dystrophy, leucoplakia, reticular skin pigmentation and pancytopenia (family DCR177). Both parents were heterozygous for one of the mutations (mother c.376G>A, father c.460T>A) and remained asymptomatic, as was a sibling with a normal genotype (figure 5.1c) (Published Vulliamy et al. 2008).

Proving evidence of a disruption in telomerase function and telomere maintenance, short telomeres are a common feature of DC and investigating this in these patients revealed very short telomeres compared to unaffected relatives and healthy controls (Work carried out by Dr Tom Vulliamy. Published Walne et al. 2007; Vulliamy et al. 2008).

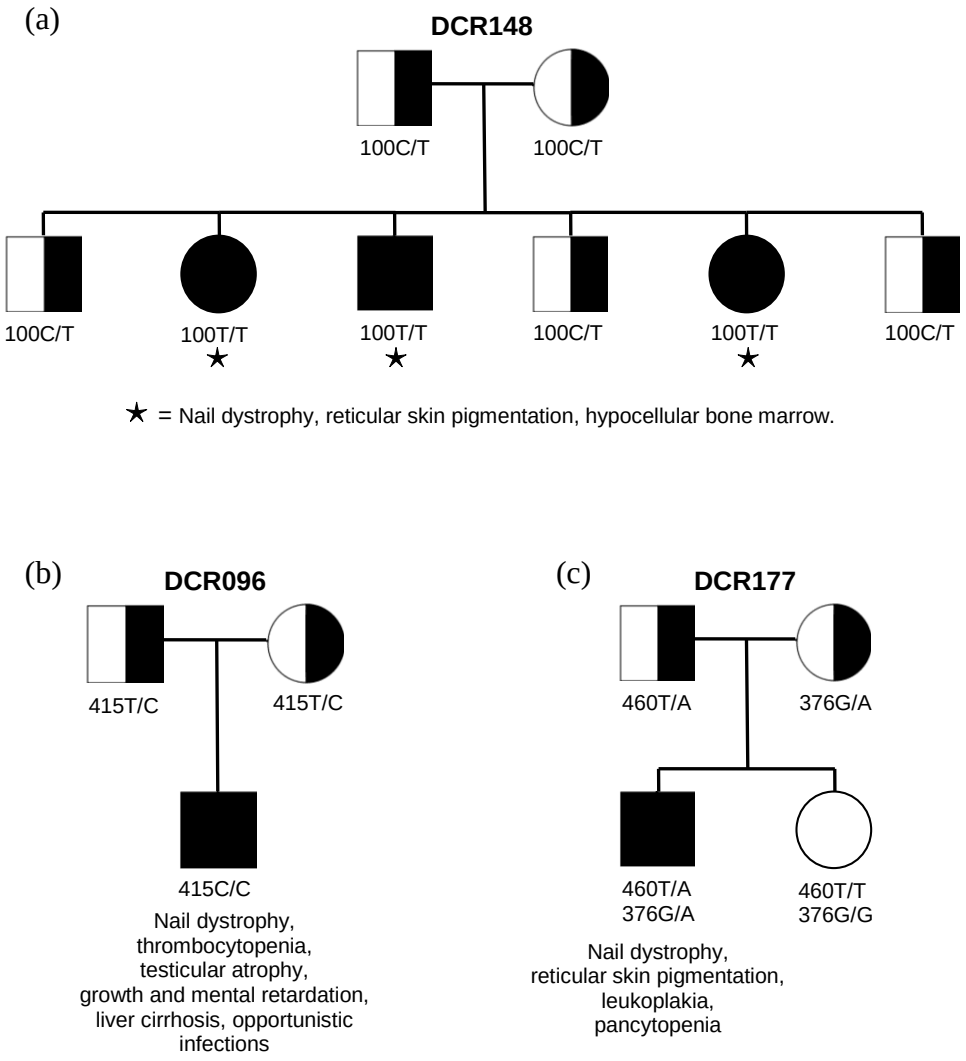


Figure 5.1. Family pedigrees. DC families with (a) *NOP10* or (b & c) *NHP2* mutations demonstrate recessive inheritance of the disease. Half filled symbols represent unaffected heterozygous and open symbols wild type family members respectively, filled symbols the affected homozygous or compound heterozygous patients with their abnormalities listed. Genotypes for the mutation are marked below each family member.

The crystal structure of an archaeal H/ACA RNP shows that dyskerin, NOP10 and NHP2 directly interact with the RNA component, stretching it across the surface of the complex (Li and Ye 2006; Hamma and Ferre-D'Amare 2010), and the RNAi experiments described in chapter 3 have demonstrated that all three of these proteins are crucial for the integrity and accumulation of TERC in human cells. Similarly, patients with *DKC1* mutations have been found to have reduced levels of TERC and experiments in mice have demonstrated that mutations affecting dyskerin directly impair TERC accumulation (Mitchell et al. 1999b; Mochizuki et al. 2004; Wang and Meier 2004). Similarly, TERC levels were found to be significantly reduced in patients with *NHP2 or NOP10* mutations, implying that these mutations also directly impair the ability to accumulate TERC (Work carried out by Yuka Masunari and Dr Amanda Walne. Published Walne et al. 2007; Vulliamy et al. 2008).

To investigate this, experiments were conducted to suppress the endogenous levels of either *NHP2* or *NOP10* and exogenously reconstitute expression with the corresponding mutant so that their impact on TERC accumulation could be evaluated.

Note: - Only the homozygous *NHP2* c.415T>C mutation was investigated in this study.

5.1.2. Experimental Procedure

HeLa cell lines stably transfected to individually express wild-type or mutant *NHP2*, wild-type or mutant *NOP10* or an empty vector control, were already available in the laboratory. PCR amplified form cDNA, the *NHP2* and *NOP10* coding sequences had been separately cloned into the expression vector pEF1/V5-His (Invitrogen) before introducing the *NHP2* c.415T>C and *NOP10* c.100C>T mutations by site directed mutagenesis to create both wild-type and mutant *NOP10* and *NHP2* expression constructs. HeLa cells had subsequently been transfected with these constructs or the empty vector under neomycin selection to create polyclonal stable cell lines for each (Work carried out by Dr Michael Kirwan).

In independent experiments, cells from these lines were transiently transfected with siRNA (*NHP2* siRNA 1 or *NOP10* siRNA 2) to reduce endogenous *NHP2* or *NOP10* levels alongside untransfected and mock transfected controls. After 48hr RNA was extracted, first strand cDNA synthesised and tested for quality and DNA contamination by PCR of the endogenous *ABL* gene (primers ABL1 and ABL2; data not shown). *NOP10* or *NHP2* expression was assessed alongside TERC levels by qPCR using TaqMan® probes (details in appendix B), calculated relative to the untransfected control for comparison (described in chapter 2.2.9). Importantly the siRNAs chosen are specific to 3' UTR regions that were not present in the corresponding expression constructs. By virtue of this, expression from these constructs was immune to the silencing effect of these siRNAs. Similarly, for *NHP2* it was possible to distinguish endogenous from total (endogenous and exogenous) expression levels by using an additional TaqMan® probe set specific to the 3'UTR (NHP2.2 – appendix B) and, therefore, only capable of detecting the endogenous transcript.

5.1.3. Results

5.1.3.1. Mutant NHP2

After 48hr the *NHP2* siRNA substantially reduced endogenous *NHP2* expression to less than 15% relative to untransfected controls in all three cell lines (figure 5.2a.). Total *NHP2* expression in the empty vector cell line was consistent with endogenous levels under siRNA treatment as expected. In contrast, total *NHP2* levels in the wild-type and mutant *NHP2* cell lines were higher, between 30-40% (figure5.2b), indicating expression from the *NHP2* constructs reconstituting that of the suppressed endogenous *NHP2* gene. Despite this rescue effect, compared to controls *NHP2* expression was limited under treatment with the siRNA and consequently corresponding TERC levels were reduced in all three cell lines (figure 5.2c). However, while TERC was reduced comparably in the empty vector and mutant *NHP2* cell lines, levels were higher in the wild-type *NHP2* cell line (46%, 49%, 61%, respectively). This suggests that the mutant NHP2 protein cannot effectively support TERC accumulation compared to the wild-type NHP2 protein.

5.1.3.2. Mutant NOP10

It was not possible to distinguish endogenous from total *NOP10* expression in these experiments but the analysis shows that after 48hr *NOP10* levels were substantially reduced compared to untransfected controls in all three lines when treated with the *NOP10* siRNA (figure 5.3a). *NOP10* levels in the wild-type and mutant cell lines signify only a very low level of expression from the *NOP10* constructs, though a less pronounced knockdown effect in these lines compared to the empty vector cell line would indicate a small level of reconstituted expression. With overall *NOP10* expression diminished under treatment with the siRNA, corresponding TERC levels were reduced in all three cell lines (figure 5.3b) but remained slightly higher in the wild-type *NOP10* cell line than in both the empty vector and mutant *NOP10* cell lines (27%, 46%, 39%, respectively). This small effect may suggest that the mutant NOP10 protein is less capable of supporting TERC accumulation than the wild-type NOP10 protein.

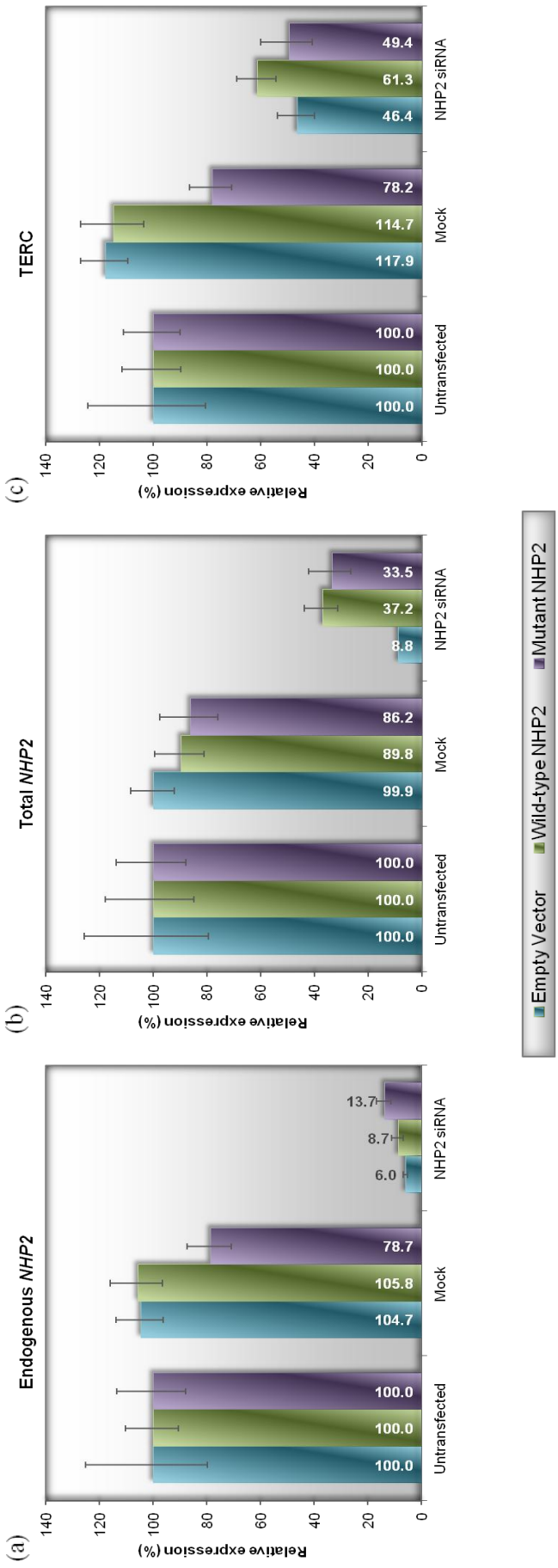


Figure 5.2. *NHP2* and TERC levels in the mutant *NHP2* experiment. (a) Endogenous *NHP2*, (b) total *NHP2* and (c) TERC levels in the empty vector control, wild type *NHP2* and mutant *NHP2* cell lines 48 hours post transfection with *NHP2* siRNA; calculated relative to the untransfected control. Values are from a single experiment, error bars marking the SD between assay triplicates.

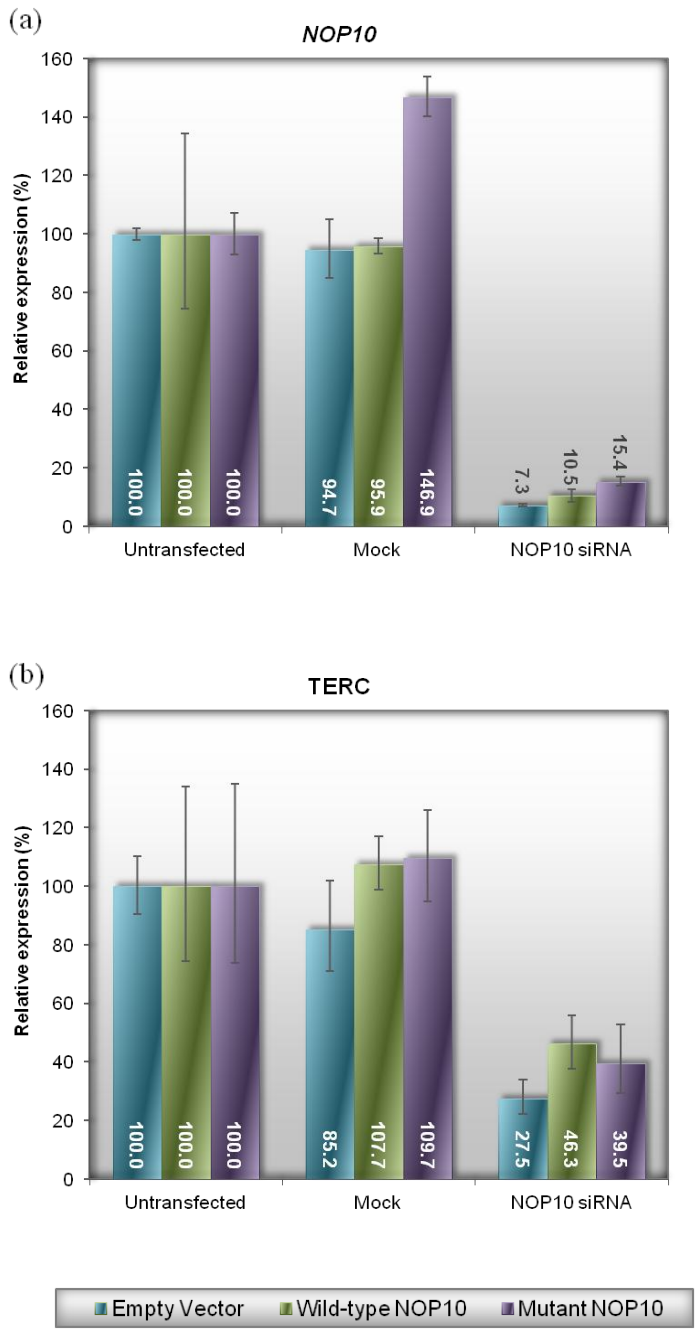


Figure 5.3. *NOP10* and TERC levels in the mutant *NOP10* experiment. (a) Total *NOP10* expression and (b) TERC levels in the empty vector control, wild type *NOP10* and mutant *NOP10* cell lines 48 hours post transfection with *NOP10* siRNA; calculated relative to the untransfected control. Values are from single experiments, error bars marking the SD between assay triplicates.

5.1.4. Discussion

The aim here was to investigate the implication that homozygous *NHP2* and *NOP10* mutations impair the accumulation of TERC. Attempting mutant reconstitution experiments, the endogenous expression of either *NHP2* or *NOP10* was successfully reduced over a 48hr period in cell lines stably expressing a corresponding wild-type or mutant construct. However the exogenous expression was limited and did not fully reconstitute that of the suppressed endogenous gene. In the *NHP2* experiment, total *NHP2* expression levels were reconstituted to between 30-40% in both the wild-type and mutant *NHP2* cell lines. While in the *NOP10* experiment, total *NOP10* levels indicated only a very low level of reconstituted expression (10-15%) in both the wild-type and mutant *NOP10* cell lines. Consequently, in both experiments corresponding TERC levels were reduced regardless of mutation status. Despite this, differences could still be seen to suggest that TERC accumulation was further impaired in both the mutant *NHP2* and *NOP10* cell lines compared to their wild-type counterparts. These results would support the implication that both the homozygous *NHP2* c.415T>C and *NOP10* c.100C>T mutations cause the reduced TERC levels in the affected patients and their disease (published Walne et al. 2007; Vulliamy et al. 2008).

Experiments described earlier in this study have demonstrated that both NOP10 and NHP2 are important to the integrity of the telomerase complex and its ability to recruit TERC. The impaired capacity to accumulate TERC demonstrated both here and in the affected patients would suggest that this integrity and recruitment is compromised by these mutations. Mapping these mutations on to the crystal structure of an archaeal H/ACA RNP offers further insight and support for this (figure 5.4). In this structure the mutated NOP10 residue R34 directly binds to the H/ACA RNA as part of a highly conserved segment of the protein (residues 34-38) that interacts with both dyskerin and the H/ACA RNA simultaneously. A substitution at this residue would be predicted to disrupt this binding as it relies on the long side chain of arginine to be inserted into the major groove of the RNA (Li and Ye 2006). The mutated NHP2 residue Y139 is located towards the C-terminal end of the protein and does not appear to be involved in the interactions with either NOP10 or the H/ACA RNA. Therefore, disruptions arising from the substitution of this residue are more difficult to predict but as a highly conserved residue it is suggested to play an important role in the protein.

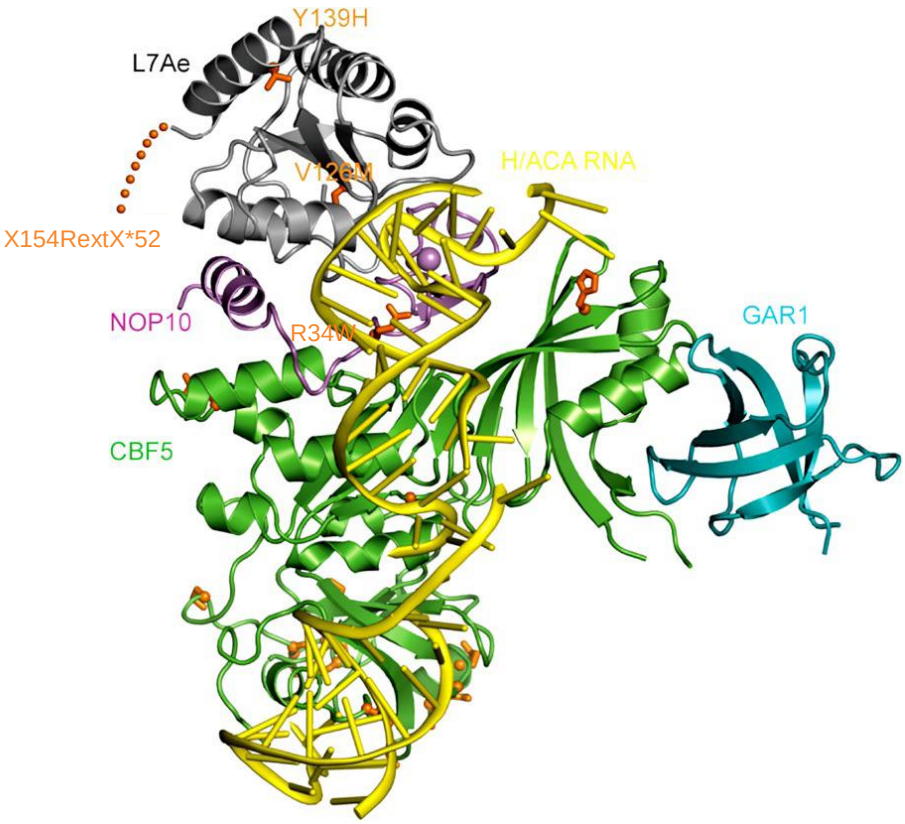


Figure 5.4. DC mutations mapped onto the crystal structure of a (Pf) H/ACA RNP. Mutated residues and side chains in NOP10 (Nop10), NHP2 (L7Ae) and dyskerin (Cbf5) are highlighted in orange with the C-terminal extension created in NHP2 also represented. The NOP10 and NHP2 mutations are labelled. (Figure adapted from Vulliamy et al. 2008, initially prepared by Kepiong Ye for Dr Vulliamy)

The disruption caused by these mutations has been more clearly demonstrated since their publication. Trahan et al used a cell free *in vitro* system to investigate the impact of DC mutations on H/ACA RNP formation, specifically protein-protein interactions and H/ACA RNA recruitment (Trahan et al. 2010). They report that the NOP10 mutation p.R34W (c.100C>T) did not disrupt formation of the protein core but its ability to bind and recruit H/ACA RNAs, including TERC, was severely impaired. The NHP2 mutation p.Y139H (c.415T>C) was found to have a more severe impact and impaired the interaction of NHP2 with NOP10, disrupting formation of the protein core and H/ACA RNA recruitment. Trahan et al also investigated the

compound heterozygous NHP2 mutations that have not been examined here. They report that p.V126M (c.376G>A) similarly disrupts formation of the protein core and H/ACA RNA recruitment but less severely than the p.Y139H mutation. Surprisingly, p.X154RextX*52 (c.460T>A) was found to have no impact on the formation of the protein core or H/ACA RNA recruitment. As recessive inheritance still implicates this mutation in the disease the authors logically suggest that its impact must be further down the biogenesis pathway. Indeed, it is important to consider the limitations of the cell free *in vitro* system used. Complex formation is contrived and simplified in this system but *in vivo* the stoichiometry of this process is likely to be more complex. It is plausible that this mutation may still have an impact through *in vivo* factors that are neutralised in this cell free context, such as chaperone interactions or protein localisation. Alternatively, the complex formed with this mutant may simply be aberrant or unstable and degraded more rapidly *in vivo*. Trahan et al have, therefore, demonstrated that H/ACA RNP integrity and H/ACA RNA recruitment are directly compromised by these mutations, providing a mechanism for the impaired accumulation of TERC seen in the affected patients and suggested in the reconstitution experiments described here.

In conclusion, these experiments sought to investigate the impact of homozygous *NOP10* and *NHP2* missense mutations observed in patients with DC. As indicated by patient TERC levels, these reconstitution experiments provide evidence to suggest that TERC accumulation is impaired as a direct consequence of these mutations. Subsequent studies by Trahan et al confirm that these mutations have a direct negative impact on H/ACA RNP integrity and H/ACA RNA recruitment, including TERC (Trahan et al. 2010). Together, these findings provide substantive evidence and a mechanism to directly incriminate these mutations in the disease of these patients; through a telomerase defect that serves to further implicate defective telomere maintenance in DC.

5.2. Dyskerin Mutations

5.2.1. Introduction

Short telomeres are a common feature of DC and the mutations identified in various telomerase components strongly implicate this complex and defective telomere maintenance in DC disease. However, sharing a common protein core with all other all H/ACA RNPs there is functional overlap that potentially implicates pseudouridylation defects in DC as well, specifically from mutations affecting dyskerin, NOP10 and NHP2.

Dyskerin mutations account for the largest proportion of characterised DC cases, most are unique with a small number seen in multiple cases and one that is particularly common (p.A353V) (Vulliamy et al. 2006). Mapping dyskerin mutations onto the archaeal crystal structure of the H/ACA RNP (refer to figure 5.4) or the proposed structural model of dyskerin (Rashid et al. 2006) (figure 5.5) demonstrates the large number (including p.A353V) that are located in or around the RNA binding PUA domain.

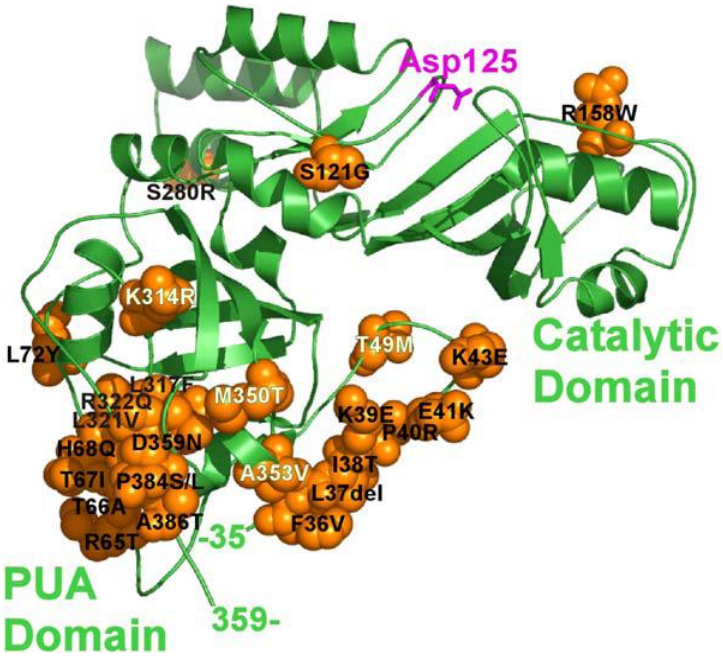


Figure 5.5. Dyskerin mutations mapped onto a model of the human protein structure. Created based on homology between dyskerin and (Pf) Cbf5. The mutated residues are highlighted in orange, white labels indicating those that occur in multiple families. (Figure taken from Rashid et al. 2006).

It is plausible to assume that these mutations may weaken or disrupt dyskerin-H/ACA RNA interactions and would therefore suggest that their role in DC is mediated through influences on H/ACA RNAs rather than a direct impact on dyskerin's catalytic ability. This is consistent with the strong implications for defects in telomerase and telomere maintenance in DC which are not directly dependant on the catalytic activity of dyskerin.

One study has investigated the impact of two dyskerin mutations, p.A353V and p.G402E, when introduced into mouse ES cells and reported marked decreases in some H/ACA RNAs and pseudouridylation for both, but only p.A353V was associated with reduced TERC levels, telomerase activity and telomere length (Mochizuki et al. 2004). On the other hand, studies of dyskerin mutations in patient material have not reached similar conclusions and found no evidence to suggest that H/ACA RNA levels or pseudouridylation are affected only TERC levels, telomerase activity and telomere length (Mitchell et al. 1999b; Wong et al. 2004). However, it is noteworthy that the mutations investigated in patient studies differ from those in the mouse study. Different again, a cell free *in vitro* study that specifically investigated the dyskerin-TERC interaction has shown that this to be severely disrupted by both the p.A353V and p.G402E mutations (Ashbridge et al. 2009). Yet, several dyskerin mutations had no impact on protein interactions or H/ACA RNA recruitment in the cell free study on H/ACA RNA assembly by Trahan et al, except for a mild decrease in TERC with the p.A353V mutation (Trahan et al. 2010). Despite this, reduced TERC levels have been reported in a large number of patients with various dyskerin mutations but other H/ACA RNAs were not investigated for comparison (Walne et al. 2007; Vulliamy et al. 2008).

Therefore, the impact of dyskerin mutations on H/ACA RNAs in DC patients remains somewhat ambiguous. To investigate this, a panel of H/ACA RNAs were quantified in primary material from DC patients with dyskerin mutations and compared to those of healthy controls.

5.2.2. Experimental Procedure

All the individuals included in this study were recruited to the DCR in London with written consent and ethical approval as defined in chapter 2.1.6. Where cDNA samples were available, 13 patients with confirmed dyskerin mutations were selected (6 x p.A353V, 3 x p.G402E, p.P40R, p.M350T, p.H68T, and p.L314Del) along with 12 healthy controls. For both sample sets, absolute levels for the H/ACA RNAs: U19, U64, U68, U70, and the C/D box RNA U97, were determined by qPCR using SYBR green, normalised to that of the *ABL* endogenous control and expressed as a target/ABL ratio for comparison (described in chapter 2.2.9, primers in appendix B). Where appropriate *P* values were calculated using the Mann-Whitney U test (described in chapter 2.2.27).

5.2.3. Results

The results of this analysis suggest that not all H/ACA RNAs are affected by dyskerin mutations (figure 5.6). Compared to healthy controls the ratios for TERC, and less strikingly, U19 and U68 are significantly lower in the DC patient group (p=0.0002, p=0.0226, p=0.0133, respectively). The other two, U64 and U70, as well as the control C/D box RNA U97, show no significant differences (p=0.4303, p=0.5316, p=0.2648, respectively).

Analysing the patient cohort by mutation shows that this discriminatory impact varies for each (figure 5.7). Further, this analysis highlights a striking consistency to the impact on TERC that is not seen with any of the other H/ACA RNAs investigated. Due to insufficient numbers statistical significance could only be determined for the p.A353V and p.G402E mutations which highlights that only TERC is significantly reduced with the p.A353V mutation (p=0.01), but both TERC and U19 reductions are significant with the p.G402E mutation (p=0.0115, p=0.0115, respectively).

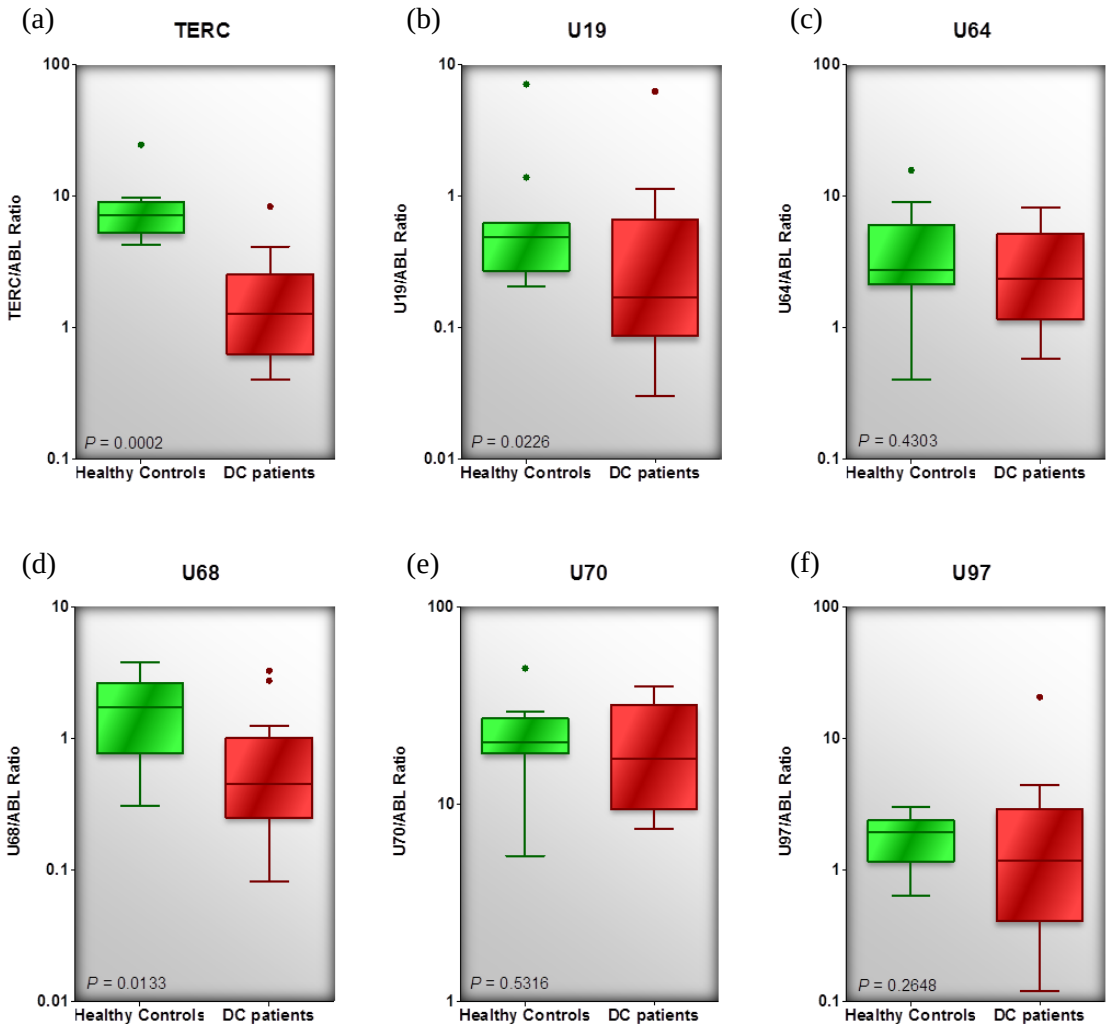


Figure 5.6. H/ACA RNA levels in patients with dyskerin mutations. Box whisker plots comparing H/ACA RNA levels in 12 healthy controls and 13 DC patients. Calculated as a target/ABL ratio the data is plotted on a log scale to correct for a skewed distribution with median values and outliers (>1.5 box lengths) marked. U97 is a C/D Box RNA control.

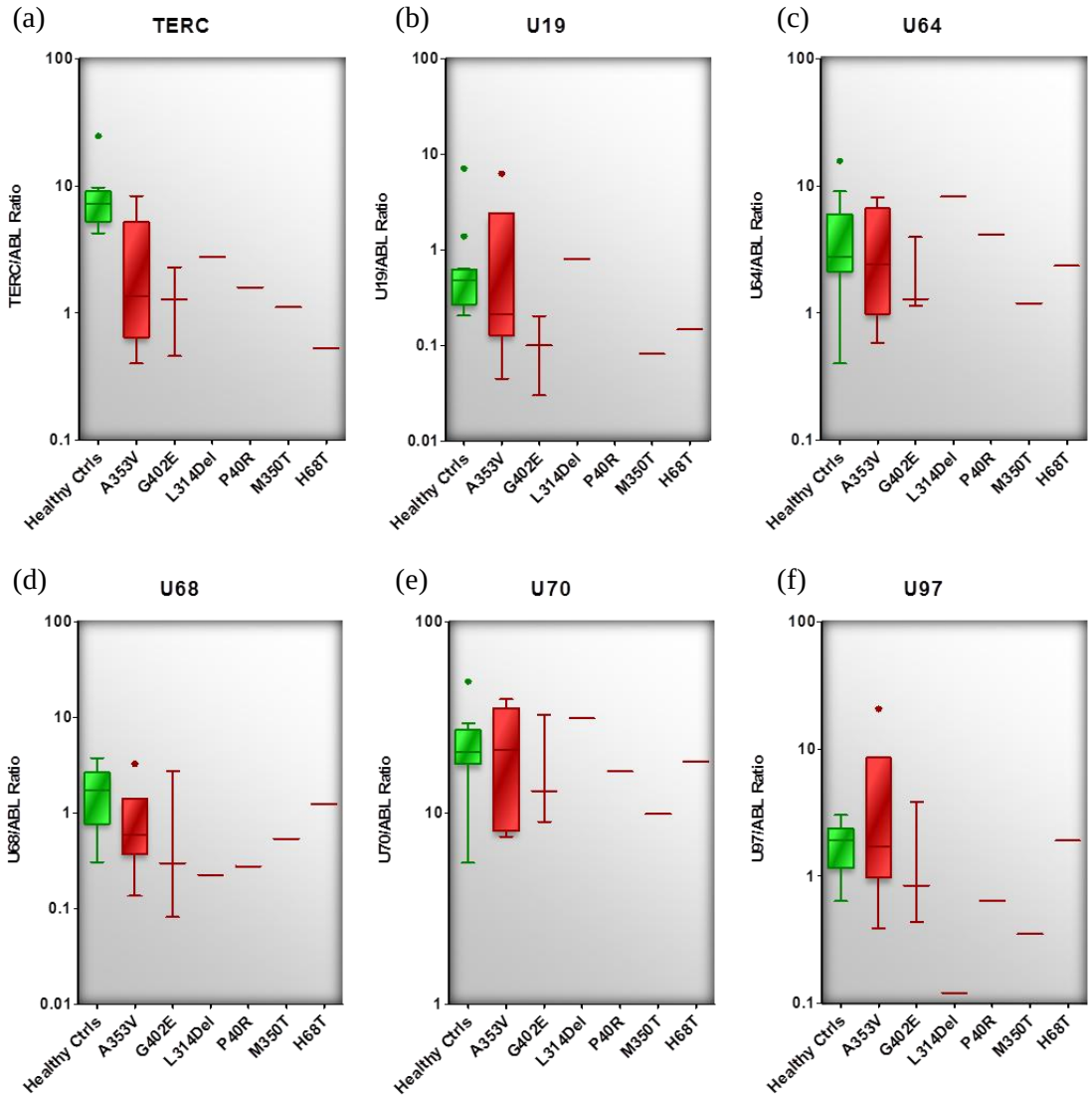


Figure 5.7. Differences in the impact on H/ACA RNA levels. Box whisker plots showing the variable impact of different dyskerin mutations on H/ACA RNA levels. Calculated as a target/ABL ratio the data is plotted on a log scale to correct for a skewed distribution with median values and outliers (>1.5 box lengths) marked.

5.2.4. Discussion

Levels of the H/ACA RNAs: TERC, U19, U64, U68, U70, and the control C/D box RNA U97, were measured in control and patient cohorts and show that while TERC is significantly reduced in the patient group some but not all of the other H/ACA RNAs are similarly affected, namely U19 and U68. Corroborating with previous patient studies to indicate that dyskerin mutations adversely affect TERC recruitment and accumulation, this observation goes further to suggests that they have a variable impact on H/ACA RNAs generally and discriminate against certain species only. Analysing each individual mutation demonstrates further variability while highlighting TERC as the only one to be affected consistently.

With the exception of TERC, there is a lack of data in the literature regarding H/ACA RNA levels in DC patients for comparison. One study investigated TERC, U17, U64 and U85 levels in primary material from a single patient with the dyskerin mutation p.Q31E and found no observable differences except decreased TERC (Wong et al. 2004). A separate study similarly found TERC to be reduced but U64, E1 and E2 unaffected in EBV transformed lymphoblast cell lines and primary fibroblasts from two different patients with dyskerin mutations, one harbouring a p.T66A mutation the other a p.L37Del (Mitchell et al. 1999b). These specific mutations are not represented in this study for direct comparison, but further demonstrate that not all H/ACA RNAs are affected by dyskerin mutations and were initially considered to suggest a selective impact on TERC alone. Taken in retrospect with the findings from the patient analysis here, it could be suggested that these studies provide further support for a consistent impact on TERC while highlighting that the restricted H/ACA RNA analysis performed was a limitation that allowed other discriminatory affects to be overlooked. A mouse ES cell study investigating the p.A353V and p.G402E mutations analysed a broader selection of H/ACA RNAs and also observed a selective impact (Mochizuki et al. 2004). They report decreases in the levels of TERC, U68, U64 and U70 for the common p.A353V mutation but only in U64 levels for the p.G402E mutation. In the patient cohort analysed here only TERC was significantly decreased with the p.A353V mutation while the p.G402E mutation significantly reduced both TERC and U19. This disparity may reflect species variation between mice and humans or variation between cases of the same mutation, as is seen in the patient cohort.

However, both studies agree in suggesting that dyskerin mutations affect multiple H/ACA RNAs in a discriminatory impact.

It has been specifically shown that the p.A353V and p.G402E mutations severely disrupt the dyskerin-TERC interaction (Ashbridge et al. 2009). However, a cell free *in vitro* study into H/ACA RNP assembly found that several dyskerin mutations had no impact on the recruitment of U17, U64, U92 and, surprisingly, TERC which was only mildly impaired with the p.A353V mutation (Trahan et al. 2010). This would imply that dyskerin mutations alone do not cause sufficient disruption to prevent H/ACA RNA recruitment to the complex as a whole, stressing the involvement of NOP10 and NHP2 in this interaction. However, this is inconsistent with the impacts observed in patients and indicates compounding influences *in vivo*. It is conceivable that the stability of the H/ACA RNA interaction is diminished in these complexes; their simplified stoichiometry and exaggerated 'turn over' *in vitro* may compensate for this and overstate the stability of the RNA recruitment compared to that *in vivo*. This potentially suggests that the impact of these mutations can be influenced by the stoichiometry of the complex and so perhaps the discrimination regarding which H/ACA RNAs are affected relates to differences in their individual expression and stability. Comparing the relative abundances of the H/ACA RNAs analysed in this study shows that the unaffected U64 and U70 are much more abundant than the affected U19 and U68, but any correlation is undermined by TERC which is highly abundant but severely affected (figure 5.8).

However, the consistency of the impact on TERC suggests that it is particularly susceptible to these mutations compared to other H/ACA RNAs. Accommodating its unique pseudoknot domain, TERC is larger (451nt against an average 100-150nt) than other H/ACA RNAs and more complex in structure, its interactions with the RNP are likely to be more complex and therefore more susceptible to disruption. Certainly, it is conceivable that its larger size alone may render this interaction more vulnerable to disruption and it is noteworthy that the study by Trahan et al, which observed normal TERC recruitment in the presence of dyskerin mutations, did not use the full length TERC molecule. Studies in yeast have suggested that the stability of the H/ACA RNA-RNP interaction differs for each H/ACA RNA (Henras et al. 2004b). Therefore, the discriminatory impact on H/ACA RNAs probably reflects combined differences in

their individual structures and specific interactions with the RNP as a whole, as well as their relative abundances.

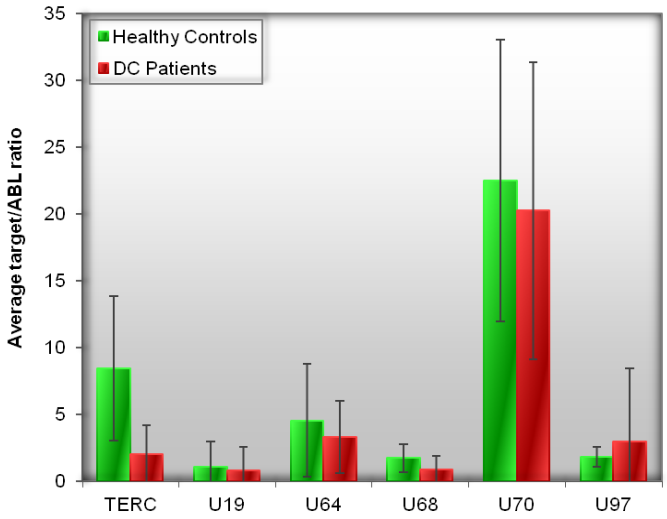


Figure 5.8. Relative abundance of the different H/ACA RNAs. Compared in both the healthy control (n=12) and DC patient (n=13) groups, averaged from target/*ABL* ratios measured in each cohort.

In conclusion, analysing different H/ACA RNAs in a cohort of DC patients has revealed that only some are significantly affected, varying for each mutation. Strikingly, the only RNA found to be consistently affected was TERC, indicating that it is highly susceptible to mutational disruptions. In fact, TERC decreases are consistently observed across various knockdown, mutation and patient studies investigating dyskerin, NOP10 or NHP2, which demonstrates that it is vulnerable to disruptions affecting the RNP complex as a whole. Combining the observations of this study with those elsewhere suggests that dyskerin mutations have a discriminatory impact on H/ACA RNAs that may relate to differences in their expression and the stability of their individual interactions with the RNP complex. Though more research is needed to elucidate the functional repercussions on pseudouridylation activity, any impact is likely to be limited and variable in comparison to the consistent impact that is observed on TERC and telomerase function - highlighting its predominant role in DC pathology.

Chapter 6. Telomere Length in Dyskeratosis Congenita.

6.1. The Prevalence of Short Telomeres in DC

6.2. Introduction

When this study began DC had been associated with mutations affecting five different genes, all of which could be linked to telomere maintenance through the telomerase complex. Since then mutations have been reported that implicate three other genes in this disease: *TINF2*, *TCAB1*, and *C16orf57* (Savage et al. 2008; Walne et al. 2010; Zhong et al. 2011). *TINF2* encodes the core shelterin component TIN2; shelterin is important for telomere stability and telomerase recruitment (de Lange 2005; Abreu et al. 2010). The specific impact of these mutations has not been confirmed but the affected patients were found to have very short telomeres (Savage et al. 2008; Walne et al. 2008). *TCAB1* encodes the newly identified telomerase component TCAB1 which is important for telomerase localisation and trafficking through cajal bodies to the telomeres (Venteicher et al. 2009). TCAB1 mutations directly impair the correct localisation of telomerase and short telomeres are observed in the affected patients (Zhong et al. 2011). Initially uncharacterised, *C16orf57* has recently been shown to encode the protein USB1, important for modifying the 3' end of the spliceosomal U6 snRNA, part of the pre-mRNA splicing machinery (Hilcenko et al. 2012; Mroczek et al. 2012). This is the first DC gene not to implicate defective telomere maintenance as the underlying cause of disease and affects only a small subset of patients, all of whom had normal length telomeres (Walne et al. 2010).

Overall, mutations affecting eight different genes have now been identified in DC: *DKC1*, *TERC*, *TERT*, *NOP10*, *NHP2*, *TCAB1*, *TINF2*, and *C16orf57*, accounting for approximately 50-60% of DCR cases. As seven of these genes are linked to telomere maintenance complexes, the predominant underlying cause of DC continues to be defective telomere maintenance and telomere shortening. Indeed, short telomeres are now considered a hallmark of this disease and have even been proposed as a diagnostic marker for DC (Alter et al. 2007). However, DC has a heterogeneous clinical presentation and disease severity is seen to differ between genetic subtypes despite the common underlying defect. Though telomere length has been investigated in each of these subtypes to indicate impacts on telomere maintenance very little comparison has been made between them. Also, approximately 40-50% of DCR

patients still have no known genetic basis and telomere length has not been widely investigated in this group. Therefore, it could be argued that the true extent of the telomere length defect in DC remains to be seen and *C16orf57* mutations have demonstrated the potential for alternative defects in this disease.

In this part of the study telomere lengths were analysed and compared in an unprecedentedly large cohort of DC patients representing the most common genetic subtypes: *DKC1*, *TINF2*, *TERC* and *TERT*, as well as those with no known genetic basis.

6.3. Experimental Procedure

All the individuals included in this study were recruited to the DCR in London with written consent and ethical approval as defined in chapter 2.1.6. Telomere lengths were analysed in 102 healthy controls and 145 DC patients: 39 with *DKC1* mutations, 19 with *TERC* mutations, 11 with *TERT* mutations, 24 with *TINF2* mutations, and 52 that were genetically uncharacterised (no mutations in known DC genes). 20 AA patients were also analysed for comparison. Measurements were taken from DNA samples using MMqPCR (described in chapter 2.2.21) and expressed as a T/S ratio normalised to that of an independent reference sample included in every sample set. The Mann-Whitney U test was used to calculate *P* values for the different comparisons (described in chapter 2.2.27).

Note: - All DNA samples were assessed for quality by agarose gel electrophoresis and only those that displayed a clean DNA band (no smearing) were accepted for analysis (data not shown).

6.4. Results

The analysis performed here shows that on average telomeres were significantly shorter in DC patients compared to healthy controls ($P<0.0001$) (figure 6.1a). In contrast, telomere length in the AA patient group did not differ significantly from the healthy controls and demonstrates that BMF is not the cause of telomere shortening in DC ($P=0.6283$) (figure 6.1a).

Corroborating with studies elsewhere, telomeres were significantly shorter in the *DKC1*, *TERC*, *TERT* and *TINF2* patient groups compared to the healthy control group ($P<0.0001$, $P<0.0001$, $P=0.0011$, and $P<0.0001$, respectively) (figure 6.1b). Plotting telomere length against age enables this comparison to be made in context to normal age related telomere shortening and emphasises that the patients had excessively short telomeres for their age (figure 6.2a). All four subgroups shared a closely overlapping range of telomere length and there were no significant differences between the *TERC*, *TERT* and *TINF2* groups. However, the *DKC1* group had a comparatively broader range of telomere length that was on average significantly longer than the *TERC* and *TINF2* groups ($P=0.0038$, $P=0.0168$, respectively).

The range of telomere length was much broader in the uncharacterised patient group and did not differ significantly from the healthy control group (figures 6.1b). Nonetheless, there was considerable overlap with the other patient groups and a subgroup with very short telomeres was evident. Plotting telomere length against age further demonstrates a number of patients that had excessively short telomeres for their age but also patients in which telomere length was normal (figure 6.2b).

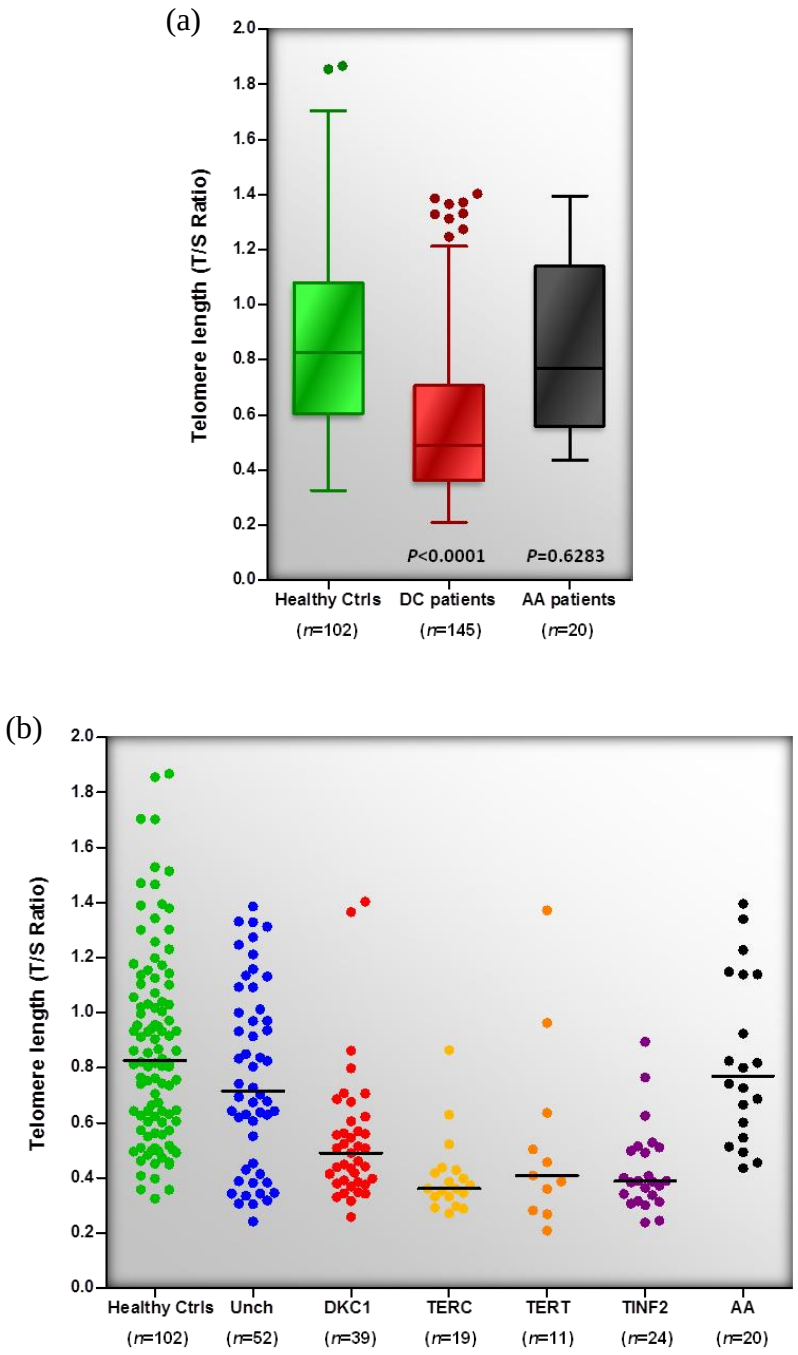


Figure 6.1. Telomere length in DC patients. (a) Box whisker plots comparing telomere length in healthy controls, DC and AA patients. (b) Comparison of telomere length in different genetic subtype of DC.

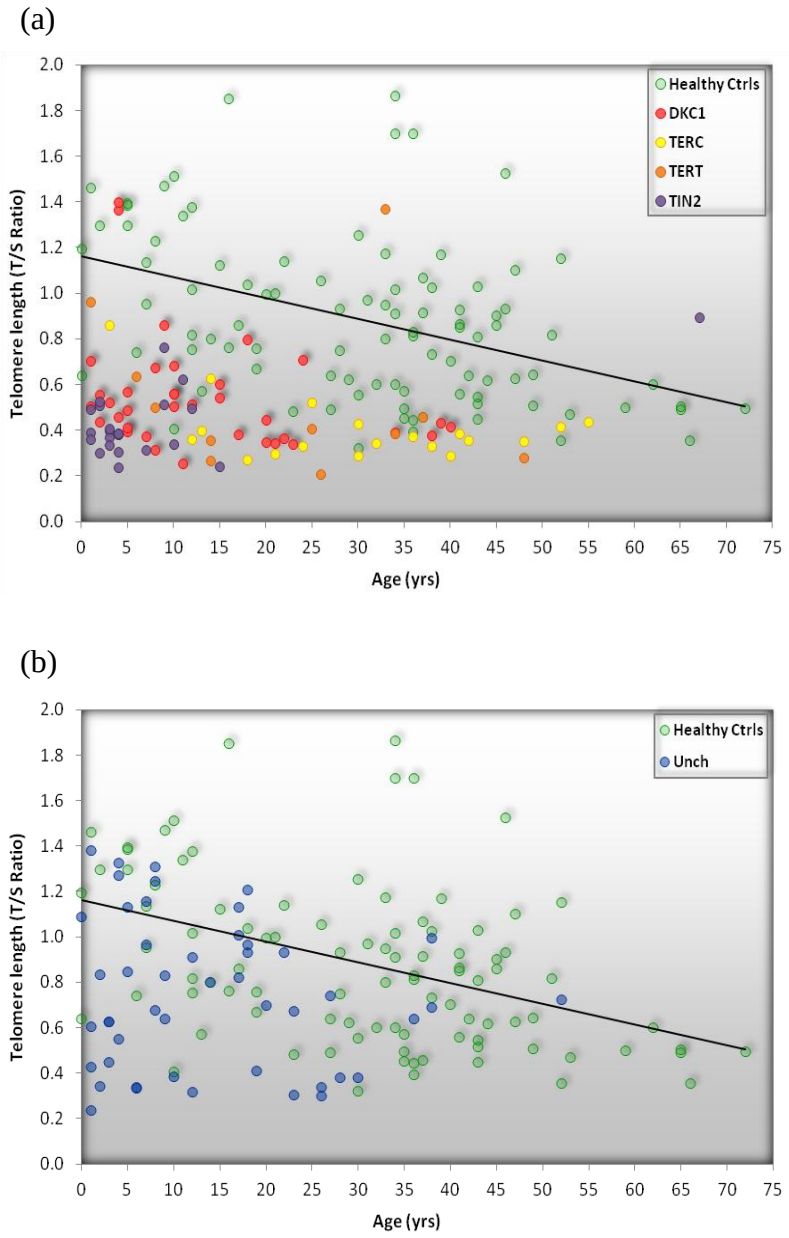


Figure 6.2. Telomere length plotted against age at report. (a) Characterised DC and (b) uncharacterised DC patients are plotted alongside the healthy controls for comparison. The trend line highlights normal age related telomere shortening in the healthy control group.

6.5. Discussion

Telomere lengths were investigated in a large cohort of 145 DC patients, differing from previous studies by encompassing a broader spectrum of DC subtypes rather than focusing on one specifically. This comprehensive analysis has reiterated that telomeres are significantly short in the four major subtypes of DC: *DKC1*, *TERC*, *TERT*, and *TINF2*, also demonstrating that telomere shortening in these patients is excessive for their age. Emphasising the prevalence of the telomere length defect in this disease, these findings are consistent with the predominant cause of DC being a defect in telomere maintenance.

DC displays a variable spectrum of clinical presentation and severity, the later generally gauged by the age of onset (particularly BMF) along with the number of somatic abnormalities: BMF before the age of 10yrs along with several other disease features classed as severe; haematological abnormalities after the age of 20yrs classed as mild (Dokal 2000). Interestingly, disease appears to manifest differently depending on the genetic subtype. *DKC1* and *TINF2* mutations are generally associated with severe disease that manifests at a very young age, *DKC1* patients mostly presenting with classical DC while *TINF2* patients often develop severe AA ahead of other DC features. In contrast, *TERC* and *TERT* mutations are typically associated with a mild disease that displays anticipation and does not manifest in all the affected generations of a family. The age of onset is more varied but typically later in life and with fewer abnormalities, predominantly presenting haematological defects.

Comparing telomere length between these subtypes reveals no correlation to their differences in disease, on average telomeres are comparatively longer in the *DKC1* group but all four subtypes share a close overlapping range of telomere length. Disease anticipation demonstrates that disease only manifests when telomere length becomes critical and so it is perhaps understandable that patients presenting disease exhibit similarly short telomeres: critically short (Vulliamy et al. 2004; Marrone et al. 2005). However, it is interesting to consider this in conjunction with differences in the typical age of disease onset for each of these groups (highlighted in figure 6.3). This indicates differences in the time taken for telomeres to become critically short and, therefore, differences in the rate of telomere shortening associated with each genetic subtype.

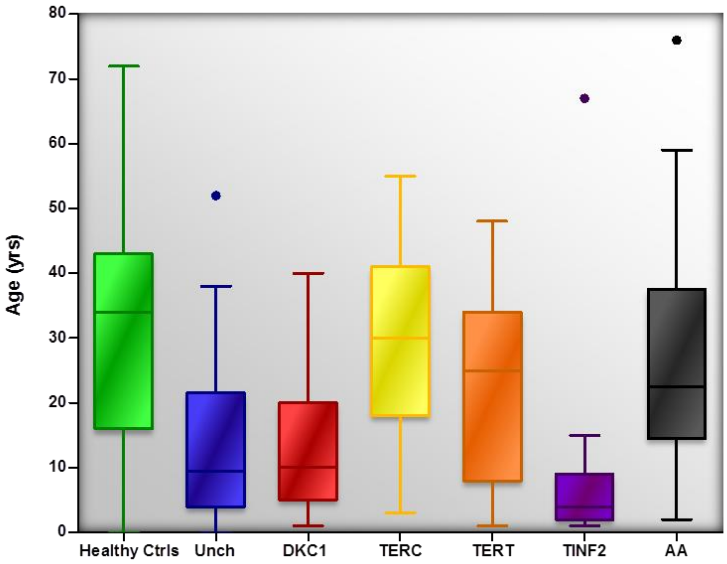


Figure 6.3. Relative age at report for patients in the different subtypes of DC.

TERC and *TERT* mutations have already been associated with a slow rate of telomere shortening that is demonstrated by the disease anticipation seen in affected families. Telomeres shorten progressively and silently through the affected generations, each inheriting shorter telomeres than the previous until after several generations critical length is reached and disease finally manifests. The age of onset and the severity of disease in these patients vary depending on the number of generations involved (Vulliamy et al. 2004; Marrone et al. 2005). This is a stark contrast to the rapid telomere shortening that is evident in the *DKC1* and *TINF2* patient groups where disease manifests in the first generation and without inheriting short telomeres. In these subtypes disease typically presents during infancy and early childhood, indicating that telomere length becomes critical within the first few years of life. Therefore, it is the rate of telomere shortening that appears to correlate with disease severity, rapid telomere shortening being associated with a severe DC phenotype.

Telomere shortening only has a pathological impact when telomeres become critically short. Directly influencing the speed with which this happens, the rate of telomere shortening governs the latency with which the telomere maintenance defect delivers a pathological impact (reflected by the age of onset). However, driven by cell division,

only cell types and tissues that are actively proliferating are susceptible to telomere shortening. Therefore, the body is more widely susceptible when in a state of rapid growth and development during early life than when it is fully developed in adulthood. Potentially, rapid telomere shortening that delivers a pathological impact during early life will have broader consequences giving rise to a more severe disease, as is evident in the *DKC1* and *TINF2* patients. Conversely, the scope of the disease arising from a latent impact that occurs in later life will be more limited and milder by comparison, as observed in the *TERC* and *TERT* patients. Therefore, by governing the latency with which the telomere length defect delivers a pathological impact, the rate of telomere shortening conceivably influences the extent and severity of the disease. This is difficult to verify empirically but supporting evidence can be found in disease anticipation where inheriting shorter telomeres brings the age of disease onset forward in successively afflicted generations and is associated with an increase in disease severity (Vulliamy et al. 2004; Marrone et al. 2005).

The range of telomere length in the *DKC1* subtype was broader and significantly longer than the other subtypes. It appears that not all of these patients have critically short telomeres, but still have disease, indicating a second defect in these patients. Other aspects of this study have provided evidence to suggest that *DKC1* mutations impact on pseudouridylation activity, though the extent of this defect has yet to be determined. However, the most severe DC phenotypes are seen in *DKC1* patients and it is intriguing that mucocutaneous features are more prominent in this group of patients compared to the *TINF2*, *TERC* and *TERT* patients. It is interesting to consider that these differences may be evidence of a contributing pseudouridylation defect in the *DKC1* patients.

Telomere length analysis in the uncharacterised group has identified a subset of patients that do not have short telomeres, indicating some other defect causing their disease. The possibility of an alternative pathology in DC has already been raised by patients with *C16orf57* mutations, which do not implicate telomere maintenance defects (Walne et al. 2010). However, perhaps it is also prudent to consider the possibility of misdiagnosis too. DC is defined and diagnosed based on clinical criteria alone but the complex spectrum of presentation and its overlap with other diseases can cause confusion. As these patients differ from the majority of DC patients at a

molecular level this raises the question: do these patients have overlapping diseases that have been confused for DC? There is argument for this with the *C16orf57* patients. Mutations in this gene were first identified in patients with poikiloderma neutropenia (PN) and have also been seen in cases of Rothmund–Thomson syndrome (RTS), two diseases that are often confused due to considerable clinical overlap that is also shared with DC (Volpi et al. 2010; Walne et al. 2010; Colombo et al. 2012). As the defect in patients with *C16orf57* mutations differs from all other characterised DC patients and in view of the clinical and genetic overlap with other diseases, some would argue that the diagnosis of DC should be reconsidered for these patients. Indeed, there is already evidence of *C16orf57* mutations being used to distinguish PN from RTS and DC (Colombo et al. 2012). As the genetic and molecular defects in uncharacterised patients with normal length telomeres are elucidated it will be interesting to see if there is further overlap with other diseases to implicate a potential misdiagnosis.

In conclusion, analysing telomere length in a large cohort of DC patients has reiterated the prevalence of short telomeres in this disease but has also demonstrated that differences in severity do not correlate with telomere length. Instead, this analysis indicates that it is differences in the rate of telomere shortening that influence the severity of this disease. However, comparatively longer telomeres in the *DKC1* patient group provides further evidence for a defect in pseudouridylation contributing to their disease that perhaps correlates with differences in the disease phenotype associated with these patients. Lastly, identifying a subset of patients that do not have short telomeres, this analysis provides further evidence of alternative defects causing DC. However, the potential for misdiagnosis also raises the question as to whether defective telomere maintenance should now be used to further define DC rather than clinical presentation alone and, therefore, whether these patients actually have DC.

6.6. Telomere Length as an Indicator of DC

6.6.1. Introduction

Touched upon in the previous section, DC has a clinical presentation that can be confused with other diseases and initial diagnosis can often be confused for AA, other DC features developing later. Similarly, a subset of AA patients that do not respond to conventional therapy have been found to have TERC mutations and short telomeres, their phenotype not sufficient to warrant a DC diagnosis but considered to be 'DC like' (Vulliamy et al. 2002; Yamaguchi et al. 2003). Correctly identifying patients with DC related disease has important implications for managing their condition though. Alter et al have reported that analysing telomere length in specific leukocyte subsets by flow FISH can be used to distinguish DC from non DC patients and propose this as a diagnostic marker in cases where there is ambiguity (Alter et al. 2007).

The DCR in receives many samples from AA patients suspected of having DC related disease, using the telomere length to further evaluate the DC potential in these patients may help to prioritize or eliminate them for DC screening. To investigate this, flow FISH was used to analyse telomere length in patients submitted to the DCR according to the criteria defined by Alter et al (Alter et al. 2007).

6.6.2. Experimental Procedure

Flow FISH (described in chapter 2.2.20) was used to determine relative telomere lengths (RTL) for lymphocyte and granulocyte subsets in blood samples obtained from 42 healthy controls and 58 patients; 16 with DC* and 42 with idiopathic AA. Measured in kilobases (kb), RTLs were calculated using a calibrator sample of predetermined length. All samples were collected with written consent and ethical approval.

* DC patients had mutations in the following genes: 3 x *TERC*, 2 x *TERT*, 2 x *DKC1*, 1 x *TINF2* and 8 were genetically uncharacterised.

6.6.3. Results and Discussion

Following the flow FISH protocol outlined by Baerlocher et al telomere lengths were successfully analysed in lymphocyte and granulocyte subsets of the healthy control and patient blood samples (Baerlocher et al. 2006). In the diagnostic criteria outlined by Alter et al, DC patients are identified as having telomeres shorter than the first percentile of healthy controls, further reporting that the lymphocyte subset provides greater sensitivity and specificity for this discrimination (Alter et al. 2007). In the analysis here 13 of the 16 DC patients had lymphocyte telomeres shorter than the first percentile (81% sensitivity) but only 10 had granulocyte telomeres below the same mark (62% sensitivity) (figure 6.4). It should be pointed out that the granulocyte subset proved variable and often few in number, making them less reliable and more difficult to analyse. Alter et al also reported difficulties analysing granulocytes in some patients (Alter et al. 2007). As the more reliable and discriminative subset, analysis of the AA patients was centred on the lymphocyte telomeres. It is also noteworthy that the 3 DC patients who did not have short lymphocyte telomeres have no known genetic basis and so it is interesting to consider that this may indicate misdiagnosis.

Of the 42 AA patients analysed 15 were found to have lymphocyte telomeres below the first percentile (figure 6.4), suggesting DC related disease and highlighting them for genetic screening. Another 5 had telomeres below the tenth percentile. Though the indicator for DC is telomeres shorter than the first percentile, this is still considered short and therefore worthy of further investigation. Subsequent genetic analysis of these 20 patients (carried out by Dr T Vulliamy) identified 3 with *TERC* mutations that confirm they have DC related AA. The rest were not found to have either *TERC* or *TINF2* mutations, the common genetic overlap between AA and DC. However, the potential for DC related disease in these patients may merit screening other DC related genes, a consideration for future projects perhaps.

Applying the criteria defined by Alter et al, using telomere length analysis by flow FISH to evaluate cases of unresponsive idiopathic AA submitted to the DCR was beneficial in identifying patients for DC screening. It has also identified 3 uncharacterised DC patients that do not have a telomere length defect, an important consideration for the future analysis and treatment of these patients. Setting up the

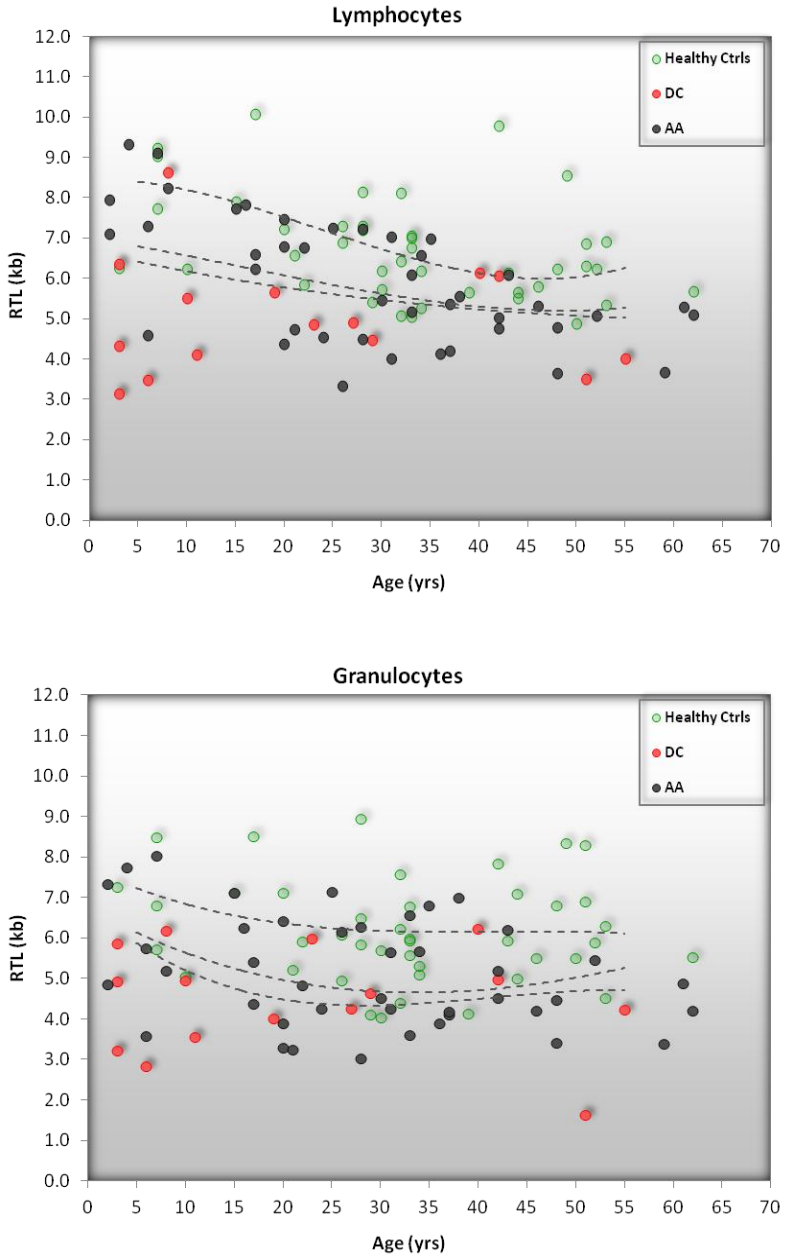


Figure 6.4. Telomere length in lymphocyte and granulocyte subsets. Plotted against age at report for DC (*n*=15) and AA (*n*=42) patients. Dashed lines plot the 1st, 10th and 50th percentiles for age groups 0-10, 10-20, 30-40, 40-50 and 50+ of the healthy controls (*n*=42).

Note: - The age at report was unknown for one DC patient, however, their lymphocyte and granulocyte telomeres were below the 1st percentile at all ages (4.3kb and 2.1kb respectively).

flow FISH procedure was relatively straight forward, however, establishing the ability to use this as a discriminatory tool proved more problematic. This requires a large number of healthy controls of evenly distributed ages to achieve well defined percentiles across the entire age range. The study by Alter et al used 400 healthy controls to accomplish this (Alter et al. 2007). Though obtaining such numbers presents a challenge in itself the bigger problem is achieving the age distribution required, especially for the very young. This study managed to obtain a cohort of only 42 healthy controls that was bias towards the 20-50 year age range. Sufficient for the purposes of this study perhaps, but in truth the normal range of telomere length is not well represented for all ages and the percentiles are badly defined. Certainly, they are a poor comparison to those used by Alter et al to define the discriminating criteria. Therefore, establishing a truly representative control cohort is highlighted as a potential difficulty for any laboratory wishing to employ this procedure and a limitation for its use diagnostically.

Chapter 7. Concluding Remarks

Prior to this study commencing, mutations affecting dyskerin, TERC and TERT had been reported in DC patients while unpublished research from our group had also identified patients harbouring mutations affecting NOP10 and NHP2. These are linked as components of the telomerase complex which is important for telomere maintenance and patients are commonly found to have short telomeres. To further examine the implications for a telomere maintenance defect in this disease and to support the publication of *NOP10* and *NHP2* as causative DC genes, this study has looked at the roles of dyskerin, NOP10 and NHP2 in the telomerase complex. Specifically, their importance, along with GAR1, for the integrity and function of telomerase has been investigated in human cells followed by *in vitro* analysis of the potential impact homozygous *NOP10* and *NHP2* missense mutations have on this complex. To consider the implications for a pseudouridylation defect, the integrity of other H/ACA RNPs has also been investigated in these experiments and in patients with *DKC1* mutations. Finally, the extent of the telomere length defect in this disease has been further investigated; telomere lengths were analysed in a patient cohort of unprecedented size and comparisons made between the different genetic subtypes of DC.

As part of telomerase; dyskerin, NOP10, NHP2 and GAR1 associate with TERC to produce an H/ACA RNP that forms the structural basis of the larger telomerase complex (reviewed Collins and Mitchell 2002). H/ACA RNP studies in yeast and mice infer that dyskerin, NOP10 and NHP2, but not GAR1, are crucial for H/ACA RNA recruitment and stability, including TERC in the context of telomerase (reviewed Meier 2005). While dyskerin has been confirmed to be important for TERC levels and telomerase activity in human cells, NOP10, NHP2 and GAR1 had not been similarly investigated (Mitchell et al. 1999b; Montanaro et al. 2006). The RNAi experiments conducted in this study confirm that in human cells TERC levels are dependent not only on dyskerin, but also NOP10 and NHP2, supporting the inference that these proteins are central to the structure of telomerase and stable TERC recruitment (Walne et al. 2007; Vulliamy and Dokal 2008). However, only dyskerin and NOP10 were found to be important for telomerase activity in these experiments. TERC is an essential component of telomerase and a limiting factor for its activity

(Marrone et al. 2004; Wong and Collins 2006). In light of its importance for stable TERC recruitment it seems unlikely that NHP2 has no bearing on telomerase activity as suggested here and it would be prudent to verify this through further investigation. Certainly, differences between the *in vitro* TRAP assay and the *in vivo* process should be taken into account. In a cell free context, incomplete but functional yeast H/ACA RNP complexes have been shown to form in the absence of Nhp2p, though the impact on H/ACA levels and pseudouridylation in Nhp2p depletion studies does not suggest their stable existence *in vivo* (Henras et al. 1998; Li et al. 2011). The possibility of similar complexes forming in the homogenised TRAP lysate (used to assay telomerase activity) should be considered. GAR1 was not found to be important for TERC accumulation or telomerase activity and is therefore not likely to be relevant for the integrity of telomerase or its activity, though again, further verification would be prudent. Also, these experiments cannot rule out the possibility of a less direct role in the function of telomerase and telomere maintenance, just as the more recently discovered TCAB1 is important for telomerase function through its role in correctly localising telomerase within the cell (Venteicher et al. 2009). As an H/ACA RNP component essential for pseudouridylation in yeast, GAR1 has now been indicated to modulate interactions with the target RNAs (Li and Ye 2006; Li et al. 2011). As this role is unnecessary in telomerase it is possible that GAR1 is simply redundant in this complex and it is noteworthy that this is the only known telomerase component not to be implicated in DC.

Though previous studies have shown that dyskerin is important for TERC levels and telomerase activity the implications on telomere length have only been investigated in mice and inferred in humans. Telomere shortening was not evident until the fourth generation in hypomorphic *DKC1* mutant mice; whereas, an *in vitro* mouse ES cell study demonstrated an immediate and steady decrease in telomere length as a result of introducing the common p.A353V mutation (Ruggero et al. 2003; Mochizuki et al. 2004). Conducting long term RNAi experiments in this study has reiterated the importance of dyskerin for sustaining TERC levels and telomerase activity over time and demonstrated that in this capacity dyskerin is imperative for maintaining telomere length in human cells. Moreover, telomeres were found to shorten rapidly when dyskerin expression was suppressed. Intriguingly, this rapid shortening stabilised and telomeres were eventually maintained at a much shorter length in the cells that

survived, also growing more slowly. It is interesting to speculate the involvement of ALT pathways here and some parallels with yeast models could be observed to suggest this, but more research is needed to verify this possibility (Henson et al. 2002; Lundblad 2002). Unfortunately, after encountering difficulties, long term experiments were limited to dyskerin only, however, similarly important for TERC levels and telomerase activity, it could be inferred that NOP10, and possibly NHP2, is likely to be necessary for maintaining telomere length as well. Together, these RNAi knockdown studies (summarised in table 2) provide support for the implications that in DC, mutations affecting either dyskerin, NOP10 or NHP2 have the potential to impact upon TERC levels, telomerase function and telomere maintenance.

Table 2. The effects of knocking down *DKC1*, *NOP10*, *NHP2* or *GAR1* expression.

Functional impact	**Knockdown experiment**			
	Dyskerin (*DKC1*)	**NOP10 (*NOP10*)**	**NHP2 (*NHP2*)**	**GAR1 (*GAR1*)**
H/ACA RNA levels	Variable reductions	Variable reductions	Variable reductions	
TERC levels	>55% decrease	>70% decrease	>60% decrease	<40% decrease
TRAP activity	>60% decrease	>70% decrease	<20% decrease	<15% decrease
Telomere length*	Shortened rapidly			
Cell growth & proliferation*	Impaired			

NOP10, NHP2 and GAR1 experiments were transient, dyskerin transient and stable. *Data from stable experiments. Blank sections indicate no data available.

In the context of DC, dyskerin mutations have already been shown to impact on TERC levels and telomerase activity (Mitchell et al. 1999b; Vulliamy et al. 2001a). Having verified that NOP10 and NHP2 are also important for the integrity of telomerase and stable TERC recruitment, reconstitution experiments were carried out to investigate the impact of homozygous missense mutations identified in these proteins. Though only low level mutant reconstitution was achieved, there was sufficient evidence to suggest that both the homozygous p.R34W NOP10 and p.Y139H NHP2 mutations directly impair TERC recruitment and accumulation. These observations offer direct support that the mutations were responsible for

reduced TERC levels in the affected patients and are disease-causing (Walne et al. 2007; Vulliamy et al. 2008). Based on the crystal structure of an archaeal H/ACA RNP, NOP10 residue R34 is directly involved in the proteins interaction with both dyskerin and the H/ACA RNA simultaneously, however, NHP2 residue Y139 is not shown to be directly involved in any of the proteins interactions within this complex (Walne et al. 2007; Vulliamy et al. 2008). Trahan et al have since reported that in a cell free system the p.Y139H NHP2 mutation prevented the protein's association with the H/ACA RNP entirely while the p.R34W NOP10 mutation severely impaired H/ACA RNA recruitment (including TERC) but did not affect formation of the protein core (Trahan et al. 2010). Together, these studies provide substantive evidence that these mutations disrupt the integrity of telomerase and TERC recruitment in the affected patients. The subsequent impact on telomerase activity has not been directly investigated in this study or elsewhere, but is inferred based on TERC levels being a limiting factor for this activity and the telomere length defects that are evident in these patients. The ambiguity surrounding NHP2's significance for telomerase activity observed in this study may detract from this argument, but the short telomeres in patients with *NHP2* mutations offers convincing support.

In the broader context, H/ACA RNPs are important for pseudouridylation, yeast and mice studies inferring that dyskerin, NOP10, NHP2 and GAR1 are all essential for this process (reviewed Meier 2005). It has since been confirmed that dyskerin (the catalytic component) is important for pseudouridylation in human cells but H/ACA RNA levels were not analysed (Montanaro et al. 2006). Investigating a selection of H/ACA RNAs, this study has observed that knocking down either dyskerin, NOP10 or NHP2 had a variable impact in which not all the RNAs were affected equally; possibly reflecting differences in their individual expression and stability. Certainly, the more stable and abundant RNAs would be favoured for recruitment into the diminished pool of H/ACA RNPs that are able to form in these experiments. Unfortunately it has not been possible to assay pseudouridylation activity directly. However, it would be reasonable to assume that the overall impact on H/ACA RNA accumulation would manifest in a corresponding degree of overall decrease in total pseudouridylation activity. These findings would support the implications for a broad impact on H/ACA RNPs in DC patients with dyskerin, NOP10 or NHP2 mutations, potentially affecting not just telomerase function but also pseudouridylation. Specific

to telomerase, TERC and TERT mutations in other patients had already indicated defective telomere maintenance to be the primary underlying cause of DC. However, of all characterised DC cases, dyskerin mutations are the most common cause and are associated with a more severe phenotype than TERC or TERT mutations; it is interesting to speculate that their disease may be exacerbated by an additional defect in pseudouridylation.

The majority of dyskerin mutations are located in the RNA binding PUA domain indicating their influence to be mediated through effects on H/ACA RNA recruitment rather than on catalytic ability directly (Rashid et al. 2006). However, the actual impact of dyskerin mutations on H/ACA RNA levels and pseudouridylation has been ambiguous. Mouse studies have demonstrated impacts that were not evident in patient studies, though it is fair to say that the later have not been extensive (Mitchell et al. 1999b; Mochizuki et al. 2004; Wong et al. 2004). In this study a broader selection of H/ACA RNAs have been analysed in a larger number of patients with more different mutations than previously reported, and it was found that dyskerin mutations have a selective impact in which only some H/ACA RNAs are significantly affected. This highlights that the analysis of only a few H/ACA RNAs is a limitation that may have prevented previous patient studies from observing any affect. Which H/ACA RNAs are impacted probably reflects differences in their specific structures and interactions with dyskerin or the RNP as a whole, as well as their relative stability and expression. Strikingly, TERC was the only RNA to be consistently affected in this analysis suggesting that it is particularly susceptible to disruptions in the RNP, possibly due to its larger size and more complex structure. Certainly, TERC decreases have been consistently observed in knockdown or mutation studies investigating dyskerin, NOP10 or NHP2, even when impacts on other H/ACA RNAs were not. It is evident from this analysis that dyskerin mutations do not have a universal impact on all pseudouridylation targets and emphasises the need to consider the repercussions here more specifically. Though the overall importance of pseudouridylation is understood, little is known about the particular significance of specific targets. Further research is needed to profile which H/ACA RNPs are affected and to investigate the downstream significance of their specific targets. Over 100 different H/ACA RNAs have been identified to date, with an expression profile that is likely to change for each cell type and with each mutation impacting differently; to investigate all of these is not a small

task. Therefore, the full extent and severity of the pseudouridylation defect remains to be elucidated, but is suggested to be limited and variable. In contrast, the impact on TERC, telomerase and telomere maintenance is consistently demonstrated, both in this study and elsewhere, supporting a more predominant role in the pathology of these patients (Mitchell et al. 1999b; Vulliamy et al. 2001a; Wong et al. 2004; Wong and Collins 2006).

Together, the RNAi experiments in this study have demonstrated the significance of dyskerin, NOP10 and NHP2 for H/ACA RNA recruitment and stability in human cells, which is in turn important for both telomerase activity and pseudouridylation. Moreover, this study has provided evidence to suggest that this function is diminished by mutations affecting these proteins in DC patients.

When this study began three genes (*DKC1*, *TERC* and *TERT*) had been associated with DC and mutations in two others (*NOP10* and *NHP2*) were being investigated. However, mutations in these genes only account for about a third of DC cases and during the course of this study mutations in other genes have been reported. The most significant of these is *TINF2* which implicated a second telomere maintenance complex and was found to be commonly mutated in DC (Savage et al. 2008; Walne et al. 2008; Vulliamy et al. 2011a). This gene encodes the TIN2 component of shelterin; a complex of proteins comprising: TIN2, TRF1, TRF2, POT1, RAP1 and TPP1, that interact to structure the telomere, important for protecting it from DNA repair mechanisms and regulating telomere elongation (de Lange 2005) Telomerase recruitment has also been linked to this complex (Abreu et al. 2010). TIN2 is the central component tethering the three DNA binding elements of the complex: TRF1, TRF2 and POT1 (de Lange 2005). The specific impact of TIN2 mutations is unknown but the affected patients have very short telomeres (Savage et al. 2008; Walne et al. 2008; Vulliamy et al. 2011a). Mutations have also been identified in *TCAB1* which encodes a newly discovered telomerase component TCAB1 (Zhong et al. 2011). Directly interacting with dyskerin and TERC, TCAB1 is important for the correct localisation and trafficking of active telomerase through cajal bodies to the telomere (Venteicher et al. 2009). TCAB1 mutations are rare and directly impair the correct localisation of telomerase; the affected patients had short telomeres (Zhong et al. 2011). Mutations have also been observed in a gene that is not linked to telomere

maintenance: *C16orf57* (Walne et al. 2010). Uncharacterised at the time, this gene has been found to encode USB1 which is indirectly implicated in pre-mRNA splicing through its role in modifying the 3' end of the spliceosomal U6 snRNA (Hilcenko et al. 2012; Mroczek et al. 2012). Telomere length was established to be normal in the small subset of patients with *C16orf57* mutation, but any impact on pre-mRNA splicing remains to be investigated (Walne et al. 2010). However, somewhat surprisingly the knockdown of USB1 in HeLa cells was not found to affect pre-mRNA splicing significantly (Mroczek et al. 2012).

Overall, approximately 50-60% of DCR patients are now characterised by mutations in one of eight genes, most commonly: *DKC1*, *TINF2*, *TERC*, and *TERT*, then more rarely: *NOP10*, *NHP2*, *TCAB1*, and *C16orf57*. Defective telomere maintenance and telomere shortening are indicated to be the cause of disease in seven of these genetic subtypes and short telomeres are now considered a typical feature of DC. Despite the common defect the severity of the disease phenotype differs between subtypes. Generally, *DKC1* and *TINF2* mutations are associated with a severe disease that presents in early life; *TERC* and *TERT* mutations with a milder disease that displays anticipation and only manifests in the later generation of affected families, typically later in life. Investigating telomere length in a large cohort of DC patients, this study has reiterated the point that telomeres are significantly short in these subtypes but found no correlation between telomere length and their differences in disease severity (Vulliamy et al. 2011b). Instead, this data indicates that it is a difference in the rate of telomere shortening that is most important in determining severity, with rapid telomere shortening associating with a more severe DC phenotype. The body is more widely susceptible to telomere shortening during early life, when in a state of rapid growth and development, than in maturity when it is fully developed. Governing the latency with which the telomere maintenance defect delivers a pathological impact (reflected by the age of onset), it is conceivable that the rate of telomere shortening determines the extent to which the body is affected and the severity of the disease. It is difficult to verify this empirically but the disease anticipation associated with TERC or TERT mutations provides convincing support. Demonstrating the latency of the impact here, disease only manifests in later generations when telomere length finally becomes critical. But more significantly, each successive generation develops disease

at an earlier age, associated with an increase in disease severity (Marrone et al. 2004; Vulliamy et al. 2004).

Comparing disease phenotypes more closely, it is intriguing that the mucocutaneous abnormalities are particularly prominent in the *DKC1* patient group but much less so in the *TINF2*, *TERC* and *TERT* patients where haematological abnormalities are foremost and the initial diagnosis may often be confused with AA. It is interesting to associate this with the potential for additional pseudouridylation defects in the *DKC1* subtype, evidence for which has been provided by other aspects of this study, though the extent and severity has yet to be verified. What's more, while all four subtypes shared a closely overlapping range of telomere length, the range was broader in the *DKC1* subtype and the average length significantly longer by comparison. It appears that not all of these patients have critically short telomeres, but still have features of disease, providing further evidence for a contributing pseudouridylation defect in these patients. Mucocutaneous abnormalities are also prominent in the patients with *C16orf57* mutations where telomere length is normal (Walne et al. 2010). The defect here is indicated to be mediated through impacts on the 3' modification of the spliceosomal U6 snRNA; it is perhaps interesting that this and other spliceosomal snRNAs are targets for pseudouridylation (Meier 2005; Hilcenko et al. 2012; Mroczek et al. 2012). Targeted by scaRNAs, it would be appealing to investigate the potential for a correlation here by profiling the impact of dyskerin mutations on these RNAs specifically.

This study has also observed a subset of genetically uncharacterised patients that do not have short telomeres, indicating an alternative pathology to their disease. Clearly the telomere length defect, though highly prevalent, is not universal to all cases of DC. This has already been highlighted by patients with mutations in *C16orf57*, so far the only gene associated with DC that does not implicate a defect in telomere maintenance (Walne et al. 2010). However, mutations in the this gene were first implicated in poikiloderma with neutropenia (PN) and have also been reported in cases of Rothmund-Thomson syndrome (RTS), the two diseases often confused due to considerable clinical overlap that is shared to a lesser extent with DC (Volpi et al. 2010; Walne et al. 2010; Colombo et al. 2012). In view of the clinical and genetic overlap here it seems prudent to consider the possibility of a confused diagnosis for

these patients. Like many disorders DC was characterised before any knowledge of its molecular basis and is defined and diagnosed on clinical criteria alone, but its complex spectrum of presentation and overlap with other diseases can cause confusion.

With research continuing to elucidate the genetic and molecular defects in this disease and related disorders it offers the opportunity to further define and differentiate them. This is already seen in DC where *TERC* mutations are used as a molecular diagnosis to distinguish the mild DC phenotype from AA. Short telomeres (measured in lymphocytes by flow FISH) have also been proposed as a marker for distinguishing DC from overlapping disorders (Alter et al. 2007). However, as this study has highlighted, the size of the control set needed to achieve accurate discrimination is a limitation for the widespread use of this approach. As the molecular basis to the disease caused by *C16orf57* mutations is distinct from that implicated in all other characterised cases of DC there is argument that the diagnosis in these patients should be reconsidered. Indeed, the use of *C16orf57* mutations as a molecular diagnosis to distinguish PN from RTS and DC has already been reported (Colombo et al. 2012). The question is whether DC should now be further defined by defective telomere maintenance rather than clinical criteria alone and, therefore, should patients with normal length telomeres now be considered to have an overlapping disease which has been confused for DC? It seems prudent to at least consider the possibility of misdiagnosis in these patients and to examine their disease more closely; perhaps there are discernible differences in their general phenotype, as seen between other DC subtypes. This issue aside, as the molecular pathways involved clearly differ, it may benefit future genetic research to analyse telomere lengths in all of the uncharacterised patients (approximately 40-50% of the DCR) and differentiate those that exhibit telomere maintenance defects from those that do not. It will be interesting to see if there is further overlap with other diseases as the molecular defects in these patients continue to be elucidated.

References

Abreu, E., E. Aritonovska, P. Reichenbach, G. Cristofari, B. Culp, R. M. Terns, J. Lingner and M. P. Terns (2010). TIN2-tethered TPP1 recruits human telomerase to telomeres in vivo. *Mol Cell Biol* **30**(12): 2971-82.

Alawi, F., P. Lin, B. Ziober and R. Patel (2011). Correlation of dyskerin expression with active proliferation independent of telomerase. *Head Neck* **33**(7): 1041-51.

Allsopp, R. C., E. Chang, M. Kashefi-Aazam, E. I. Rogaev, M. A. Piatyszek, J. W. Shay and C. B. Harley (1995). Telomere shortening is associated with cell division in vitro and in vivo. *Exp Cell Res* **220**(1): 194-200.

Alter, B. P., G. M. Baerlocher, S. A. Savage, S. J. Chanock, B. B. Weksler, J. P. Willner, J. A. Peters, N. Giri and P. M. Lansdorp (2007). Very short telomere length by flow fluorescence in situ hybridization identifies patients with dyskeratosis congenita. *Blood* **110**(5): 1439-47.

Alter, B. P., N. Giri, S. A. Savage and P. S. Rosenberg (2009). Cancer in dyskeratosis congenita. *Blood* **113**(26): 6549-57.

Amarzguioui, M., J. J. Rossi and D. Kim (2005). Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Lett* **579**(26): 5974-81.

Aravind, L. and E. V. Koonin (1999). Novel predicted RNA-binding domains associated with the translation machinery. *J Mol Evol* **48**(3): 291-302.

Armanios, M., J. L. Chen, Y. P. Chang, R. A. Brodsky, A. Hawkins, C. A. Griffin, J. R. Eshleman, A. R. Cohen, A. Chakravarti, A. Hamosh, et al. (2005). Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci USA* **102**(44): 15960-15964.

Ashbridge, B., A. Orte, J. A. Yeoman, M. Kirwan, T. Vulliamy, I. Dokal, D. Klenerman and S. Balasubramanian (2009). Single-molecule analysis of the human telomerase RNA.dyskerin interaction and the effect of dyskeratosis congenita mutations. *Biochemistry* **48**(46): 10858-65.

Baerlocher, G. M., I. Vulto, G. de Jong and P. M. Lansdorp (2006). Flow cytometry and FISH to measure the average length of telomeres (flow FISH). *Nat Protoc* **1**(5): 2365-76.

Bantounas, I., L. A. Phylactou and J. B. Uney (2004). RNA interference and the use of small interfering RNA to study gene function in mammalian systems. *J Mol Endocrinol* **33**(3): 545-57.

Birmingham, A., E. M. Anderson, A. Reynolds, D. Ilsley-Tyree, D. Leake, Y. Fedorov, S. Baskerville, E. Maksimova, K. Robinson, J. Karpilow, et al. (2006). 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods* **3**(3): 199-204.

Blackburn, E. H. (2001). Switching and signaling at the telomere. *Cell* **106**(6): 661-673.

Bousquet-Antonelli, C., Y. Henry, J. P. G'Elugne, M. Caizergues-Ferrer and T. Kiss (1997). A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J* **16**(15): 4770-4776.

Brown, W. R., P. J. MacKinnon, A. Villasante, N. Spurr, V. J. Buckle and M. J. Dobson (1990). Structure and polymorphism of human telomere-associated DNA. *Cell* **63**(1): 119-32.

Brummelkamp, T. R., R. Bernards and R. Agami (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**(5567): 550-3.

Castanotto, D., H. Li and J. J. Rossi (2002). Functional siRNA expression from transfected PCR products. *RNA* **8**(11): 1454-60.

Cawthon, R. M. (2009). Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res* **37**(3): e21.

Chen, J. L. and C. W. Greider (2004). Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci* **29**(4): 183-192.

Collins, K. and J. R. Mitchell (2002). Telomerase in the human organism. *Oncogene* **21**(4): 564-579.

Colombo, E. A., J. F. Bazan, G. Negri, C. Gervasini, N. H. Elcioglu, D. Yucelten, I. Altunay, U. Cetincelik, A. Teti, A. Del Fattore, et al. (2012). Novel C16orf57 mutations in patients with Poikiloderma with Neutropenia: bioinformatic analysis of the protein and predicted effects of all reported mutations. *Orphanet J Rare Dis* **7**: 7.

Cong, Y. S., W. E. Wright and J. W. Shay (2002). Human telomerase and its regulation. *Microbiol Mol Biol Rev* **66**(3): 407-425.

Darzacq, X., N. Kittur, S. Roy, Y. Shav-Tal, R. H. Singer and U. T. Meier (2006). Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* **173**(2): 207-18.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**(18): 2100-2110.

Dez, C., A. Henras, B. Faucon, D. L. Lafontaine, M. Caizergues-Ferrer and Y. Henry (2001). Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p. *Nucleic Acids Res* **29**(3): 598-603.

Dokal, I. (2000). Dyskeratosis congenita in all its forms. *Br J Haematol* **110**(4): 768-779.

Dokal, I. (2001). Dyskeratosis congenita. A disease of premature ageing. *Lancet* **358 Suppl**: S27.

Drachtman, R. A. and B. P. Alter (1995). Dyskeratosis congenita. *Dermatol Clin* **13**(1): 33-9.

Dragon, F., V. Pogacic and W. Filipowicz (2000). In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol* **20**(9): 3037-3048.

Fasching, C. L., K. Bower and R. R. Reddel (2005). Telomerase-independent telomere length maintenance in the absence of alternative lengthening of telomeres-associated promyelocytic leukemia bodies. *Cancer Res* **65**(7): 2722-9.

Fedorov, Y., E. M. Anderson, A. Birmingham, A. Reynolds, J. Karpilow, K. Robinson, D. Leake, W. S. Marshall and A. Khvorova (2006). Off-target effects by siRNA can induce toxic phenotype. *RNA* **12**(7): 1188-96.

Filipowicz, W. and V. Pogacic (2002). Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol* **14**(3): 319-327.

Fu, D. and K. Collins (2003). Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell* **11**(5): 1361-1372.

Ge, J., S. D. Crosby, M. E. Heinz, M. Bessler and P. J. Mason (2010). SnoRNA microarray analysis reveals changes in H/ACA and C/D RNA levels caused by dyskerin ablation in mouse liver. *Biochem J* **429**(1): 33-41.

Girard, J. P., H. Lehtonen, M. Caizergues-Ferrer, F. Amalric, D. Tollervey and B. Lapeyre (1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *Embo J* **11**(2): 673-82.

Gou, D., N. Jin and L. Liu (2003). Gene silencing in mammalian cells by PCR-based short hairpin RNA. *FEBS Lett* **548**(1-3): 113-8.

Greider, C. W. (1998). Telomerase activity, cell proliferation, and cancer. *Proc Natl Acad Sci USA* **95**(1): 90-2.

Gu, B. W., M. Bessler and P. J. Mason (2008). A pathogenic dyskerin mutation impairs proliferation and activates a DNA damage response independent of telomere length in mice. *Proc Natl Acad Sci USA* **105**(29): 10173-8.

Hamma, T. and A. R. Ferre-D'Amare (2010). The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* **285**(2): 805-9.

Harley, C. B., A. B. Futcher and C. W. Greider (1990). Telomeres Shorten During Aging of Human Fibroblasts. *Nature* **345**(6274): 458-460.

Hayflick, L. (1965). The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**: 614-36.

Heiss, N. S., S. W. Knight, T. J. Vulliamy, S. M. Klauck, S. Wiemann, P. J. Mason, A. Poustka and I. Dokal (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19**(1): 32-38.

Henras, A., Y. Henry, C. Bousquet-Antonelli, J. Noaillac-Depeyre, J. P. Gelugne and M. Caizergues-Ferrer (1998). Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J* **17**(23): 7078-7090.

Henras, A. K., C. Dez and Y. Henry (2004a). RNA structure and function in C/D and H/ACA s(no)RNPs. *Curr Opin Struct Biol* **14**(3): 335-43.

Henras, A. K., R. Capeyrou, Y. Henry and M. Caizergues-Ferrer (2004b). Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. *RNA* **10**(11): 1704-12.

Henson, J. D., A. A. Neumann, T. R. Yeager and R. R. Reddel (2002). Alternative lengthening of telomeres in mammalian cells. *Oncogene* **21**(4): 598-610.

Hilcenko, C., P. J. Simpson, A. J. Finch, F. R. Bowler, M. J. Churcher, L. Jin, L. C. Packman, A. Shlien, P. Campbell, M. Kirwan, et al. (2012). Aberrant 3' oligoadenylation of spliceosomal U6 small nuclear RNA in poikiloderma with neutropenia. *Blood*: Prepublished online.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**(6): 635-7.

Jackson, A. L. and P. S. Linsley (2004). Noise amidst the silence: off-target effects of siRNAs? *Trends Genet* **20**(11): 521-4.

Jackson, A. L., J. Burchard, J. Schelter, B. N. Chau, M. Cleary, L. Lim and P. S. Linsley (2006). Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **12**(7): 1179-87.

Kirwan, M., R. Beswick, A. J. Walne, U. Hossain, C. Casimir, T. Vulliamy and I. Dokal (2011). Dyskeratosis congenita and the DNA damage response. *Br J Haematol* **153**(5): 634-43.

Kittur, N., X. Darzacq, S. Roy, R. H. Singer and U. T. Meier (2006). Dynamic association and localization of human H/ACA RNP proteins. *RNA* **12**(12): 2057-2062.

Knight, S. W., T. J. Vulliamy, N. S. Heiss, G. Matthijs, K. Devriendt, J. M. Connor, M. D'Urso, A. Poustka, P. J. Mason and I. Dokal (1998). 1.4 Mb candidate gene region for X linked dyskeratosis congenita defined by combined haplotype and X chromosome inactivation analysis. *J Med Genet* **35**(12): 993-996.

Knight, S. W., N. S. Heiss, T. J. Vulliamy, C. M. Aalfs, C. McMahon, P. Richmond, A. Jones, R. C. Hennekam, A. Poustka, P. J. Mason, et al. (1999). Unexplained aplastic anaemia, immunodeficiency, and cerebellar hypoplasia (Hoyeraal-Hreidarsson syndrome) due to mutations in the dyskeratosis congenita gene, DKC1. *Br J Haematol* **107**(2): 335-339.

Lafontaine, D. L., C. Bousquet-Antonelli, Y. Henry, M. Caizergues-Ferrer and D. Tollervey (1998). The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev* **12**(4): 527-537.

Leung, R. K. and P. A. Whittaker (2005). RNA interference: from gene silencing to gene-specific therapeutics. *Pharmacol Ther* **107**(2): 222-39.

Li, L. and K. Q. Ye (2006). Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **443**(7109): 302-307.

Li, S., J. Duan, D. Li, B. Yang, M. Dong and K. Ye (2011). Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* **25**(22): 2409-21.

Lin, X., X. Ruan, M. G. Anderson, J. A. McDowell, P. E. Kroeger, S. W. Fesik and Y. Shen (2005). siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Res* **33**(14): 4527-35.

Lundblad, V. (2002). Telomere maintenance without telomerase. *Oncogene* **21**(4): 522-31.

Luzzatto, L. and A. Karadimitris (1998). Dyskeratosis and ribosomal rebellion. *Nat Genet* **19**(1): 6-7.

Marciniak, R. A., D. Cavazos, R. Montellano, Q. Chen, L. Guarente and F. B. Johnson (2005). A novel telomere structure in a human alternative lengthening of telomeres cell line. *Cancer Res* **65**(7): 2730-7.

Marrone, A. and P. J. Mason (2003). Dyskeratosis congenita. *Cell Mol Life Sci* **60**(3): 507-517.

Marrone, A., D. Stevens, T. Vulliamy, I. Dokal and P. J. Mason (2004). Heterozygous telomerase RNA mutations found in dyskeratosis congenita and aplastic anemia reduce telomerase activity via haploinsufficiency. *Blood* **104**(13): 3936-3942.

Marrone, A., A. Walne and I. Dokal (2005). Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr Opin Genet Dev* **15**(3): 249-257.

Marrone, A. and I. Dokal (2006). Dyskeratosis congenita: A disorder of telomerase deficiency and its relationship to other diseases. *Expert Rev Dermatol* **1**(3): 463-480.

Marrone, A., A. Walne, H. Tamary, Y. Masunari, M. Kirwan, R. Beswick, T. Vulliamy and I. Dokal (2007). Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal-Hreidarsson syndrome. *Blood* **110**(13): 4198-205.

Meier, U. T. (2005). The many facets of H/ACA ribonucleoproteins. *Chromosoma* **114**(1): 1-14.

Mitchell, J. R., J. Cheng and K. Collins (1999a). A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol Cell Biol* **19**(1): 567-576.

Mitchell, J. R., E. Wood and K. Collins (1999b). A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* **402**(6761): 551-555.

Mochizuki, Y., J. He, S. Kulkarni, M. Bessler and P. J. Mason (2004). Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proc Natl Acad Sci USA* **101**(29): 10756-10761.

Montanaro, L., A. Chilla, D. Trere, A. Pession, M. Govoni, P. L. Tazzari and M. Derenzini (2002). Increased Mortality Rate and Not Impaired Ribosomal Biogenesis is Responsible for Proliferative Defect in Dyskeratosis Congenita Cell Lines. *J Invest Dermatol* **118**(1): 193-198.

Montanaro, L., M. Brigotti, J. Clohessy, S. Barbieri, C. Ceccarelli, D. Santini, M. Taffurelli, M. Calienni, J. Teruya-Feldstein, D. Trere, et al. (2006). Dyskerin expression influences the level of ribosomal RNA pseudo-uridylation and telomerase RNA component in human breast cancer. *J Pathol* **210**(1): 10-8.

Montanaro, L., M. Calienni, C. Ceccarelli, D. Santini, M. Taffurelli, S. Pileri, D. Trere and M. Derenzini (2008). Relationship between dyskerin expression and telomerase activity in human breast cancer. *Cell Oncol* **30**(6): 483-90.

Montanaro, L., M. Calienni, S. Bertoni, L. Rocchi, P. Sansone, G. Storci, D. Santini, C. Ceccarelli, M. Taffurelli, D. Carnicelli, et al. (2010). Novel dyskerin-mediated mechanism of p53 inactivation through defective mRNA translation. *Cancer Res* **70**(11): 4767-77.

Morrish, T. A. and C. W. Greider (2009). Short telomeres initiate telomere recombination in primary and tumor cells. *PLoS Genet* **5**(1): e1000357.

Mroczek, S., J. Krwawicz, J. Kutner, M. Lazniewski, I. Kucinski, K. Ginalski and A. Dziembowski (2012). C16orf57, a gene mutated in poikiloderma with neutropenia, encodes a putative phosphodiesterase responsible for the U6 snRNA 3' end modification. *Genes Dev* **26**(17): 1911-25.

Nabetani, A. and F. Ishikawa (2011). Alternative lengthening of telomeres pathway: recombination-mediated telomere maintenance mechanism in human cells. *J Biochem* **149**(1): 5-14.

Pogacic, V., F. Dragon and W. Filipowicz (2000). Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Mol Cell Biol* **20**(23): 9028-9040.

Rashid, R., B. Liang, D. L. Baker, O. A. Youssef, Y. He, K. Phipps, R. M. Terns, M. P. Terns and H. Li (2006). Crystal Structure of a Cbf5-Nop10-Gar1 Complex and Implications in RNA-Guided Pseudouridylation and Dyskeratosis Congenita. *Mol Cell* **21**(2): 249-260.

Ruggero, D., S. Grisendi, F. Piazza, E. Rego, F. Mari, P. H. Rao, C. Cordon-Cardo and P. P. Pandolfi (2003). Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **299**(5604): 259-262.

Savage, S. A., N. Giri, G. M. Baerlocher, N. Orr, P. M. Lansdorp and B. P. Alter (2008). TINF2, a component of the shelterin telomere protection complex, is mutated in dyskeratosis congenita. *Am J Hum Genet* **82**(2): 501-509.

Shay, J. W. and W. E. Wright (2005). Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* **26**(5): 867-74.

Sieron, P., C. Hader, J. Hatina, R. Engers, A. Wlazlinski, M. Muller and W. A. Schulz (2009). DKC1 overexpression associated with prostate cancer progression. *Br J Cancer* **101**(8): 1410-6.

Sledz, C. A., M. Holko, M. J. de Veer, R. H. Silverman and B. R. Williams (2003). Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**(9): 834-9.

Trahan, C., C. Martel and F. Dragon (2010). Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* **19**(5): 825-36.

Tschaharganeh, D., V. Ehemann, T. Nussbaum, P. Schirmacher and K. Breuhahn (2007). Non-specific effects of siRNAs on tumor cells with implications on therapeutic applicability using RNA interference. *Pathol Oncol Res* **13**(2): 84-90.

Vaziri, H., W. Dragowska, R. C. Allsopp, T. E. Thomas, C. B. Harley and P. M. Lansdorp (1994). Evidence for a mitotic clock in human hematopoietic stem cells: loss of telomeric DNA with age. *Proc Natl Acad Sci USA* **91**(21): 9857-60.

Venteicher, A. S., E. B. Abreu, Z. Meng, K. E. McCann, R. M. Terns, T. D. Veenstra, M. P. Terns and S. E. Artandi (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* **323**(5914): 644-8.

Volpi, L., G. Roversi, E. A. Colombo, N. Leijsten, D. Concolino, A. Calabria, M. A. Mencarelli, M. Fimiani, F. Macciardi, R. Pfundt, et al. (2010). Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. *Am J Hum Genet* **86**(1): 72-6.

Vulliamy, T., A. Marrone, F. Goldman, A. Dearlove, M. Bessler, P. J. Mason and I. Dokal (2001b). The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**(6854): 432-435.

Vulliamy, T., A. Marrone, I. Dokal and P. J. Mason (2002). Association between aplastic anaemia and mutations in telomerase RNA. *Lancet* **359**(9324): 2168-2170.

Vulliamy, T., A. Marrone, R. Szydlo, A. Walne, P. J. Mason and I. Dokal (2004). Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in TERC. *Nat Genet* **36**(5): 447-449.

Vulliamy, T., R. Beswick, M. J. Kirwan, U. Hossain, A. J. Walne and I. Dokal (2011a). Telomere length measurement can distinguish pathogenic from non-pathogenic variants in the shelterin component, TIN2. *Clin Genet* **81**(1): 76-81.

Vulliamy, T. J., S. W. Knight, P. J. Mason and I. Dokal (2001a). Very short telomeres in the peripheral blood of patients with X-linked and autosomal dyskeratosis congenita. *Blood Cells Mol Dis* **27**(2): 353-357.

Vulliamy, T. J., A. Walne, A. Baskaradas, P. J. Mason, A. Marrone and I. Dokal (2005). Mutations in the reverse transcriptase component of telomerase (TERT) in patients with bone marrow failure. *Blood Cells Mol Dis* **34**(3): 257-263.

Vulliamy, T. J., A. Marrone, S. W. Knight, A. Walne, P. J. Mason and I. Dokal (2006). Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. *Blood* **107**(7): 2680-2685.

Vulliamy, T. J., R. Beswick, M. Kirwan, A. Marrone, M. Digweed, A. Walne and I. Dokal (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc Natl Acad Sci USA* **105**(23): 8073-8.

Vulliamy, T. J. and I. Dokal (2008). Dyskeratosis congenita: The diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**(1): 122-130.

Vulliamy, T. J., M. J. Kirwan, R. Beswick, U. Hossain, C. Baqai, A. Ratcliffe, J. Marsh, A. Walne and I. Dokal (2011b). Differences in disease severity but similar telomere lengths in genetic subgroups of patients with telomerase and shelterin mutations. *PloS one* **6**(9): e24383.

Walne, A. J., A. Marrone and I. Dokal (2005). Dyskeratosis congenita: a disorder of defective telomere maintenance? *Int J Hematol* **82**(3): 184-189.

Walne, A. J., T. Vulliamy, A. Marrone, R. Beswick, M. Kirwan, Y. Masunari, F.-h. Al-Qurashi, M. Aljurf and I. Dokal (2007). Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Hum Mol Genet* **16**(13): 1619-1629.

Walne, A. J., T. J. Vulliamy, R. Beswick, M. Kirwan and I. Dokal (2008). TINF2 mutations result in very short telomeres: Analysis of a large cohort of patients with dyskeratosis congenita and related bone marrow failure syndromes. *Blood* **112**(9): 3594-600.

Walne, A. J., T. Vulliamy, R. Beswick, M. Kirwan and I. Dokal (2010). Mutations in C16orf57 and normal-length telomeres unify a subset of patients with dyskeratosis congenita, poikiloderma with neutropenia and Rothmund-Thomson syndrome. *Hum Mol Genet* **19**(22): 4453-61.

Wang, C., C. C. Query and U. T. Meier (2002). Immunopurified small nucleolar ribonucleoprotein particles pseudouridylate rRNA independently of their association with phosphorylated Nopp140. *Mol Cell Biol* **22**(24): 8457-66.

Wang, C. and U. T. Meier (2004). Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J* **23**(8): 1857-1867.

Watkins, N. J., A. Gottschalk, G. Neubauer, B. Kastner, P. Fabrizio, M. Mann and R. Luhrmann (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* **4**(12): 1549-68.

Wong, J. M., M. J. Kyasa, L. Hutchins and K. Collins (2004). Telomerase RNA deficiency in peripheral blood mononuclear cells in X-linked dyskeratosis congenita. *Hum Genet* **115**(5): 448-55.

Wong, J. M. Y. and K. Collins (2006). Telomerase RNA level limits telomere maintenance in X-linked dyskeratosis congenita. *Genes Dev* **20**(20): 2848-2858.

Yaghmai, R., A. Kimyai-Asadi, K. Rostamiani, N. S. Heiss, A. Poustka, W. Eyaid, J. Bodurtha, H. C. Nousari, A. Hamosh and A. Metzenberg (2000). Overlap of dyskeratosis congenita with the Hoyeraal-Hreidarsson syndrome. *J Pediatr* **136**(3): 390-393.

Yamaguchi, H., G. M. Baerlocher, P. M. Lansdorp, S. J. Chanock, O. Nunez, E. Sloand and N. S. Young (2003). Mutations of the human telomerase RNA gene (TERC) in aplastic anemia and myelodysplastic syndrome. *Blood* **102**(3): 916-918.

Yamaguchi, H., R. T. Calado, H. Ly, S. Kajigaya, G. M. Baerlocher, S. J. Chanock, P. M. Lansdorp and N. S. Young (2005). Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. *N Engl J Med* **352**(14): 1413-1424.

Zeng, Y., R. Yi and B. R. Cullen (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc Natl Acad Sci USA* **100**(17): 9779-84.

Zhong, F., S. A. Savage, M. Shkreli, N. Giri, L. Jessop, T. Myers, R. Chen, B. P. Alter and S. E. Artandi (2011). Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev* **25**(1): 11-6.

Appendix A – siRNA Details

siRNA	5' to 3' Sequence (Sense then antisense)
NOP10 siRNA I	GGGUCCCUUAAACUGAUGUtt ACAUCAGUUUAAGGGACCCtc
NOP10 siRNA II	GGGAACACAUUUGACUUUUtt AAAAGUCAAAUGUGUUCCCtt
NOP10 siRNA III	CCCAUAAAGGGAACACAUUtt AAUGUGUUCCCUUUAUGGGtg
NHP2 siRNA I	GGUAAUGCUAGUCUUCUGUtt ACAGAAGACUAGCAUUACCtt
NHP2 siRNA II	GGCACAGAGAUCUGUAAUUtt AAUUACAGAUCUCUGUGCCtt
NHP2 siRNA III	GGCUGUUCCAAGAAUGUGGtt CCACAUUCUUGGAACAGCCtt
GAR1 siRNA I	GGACCUCCAGAACHUGUAGtt CUACACGUUCUGGAGGUCCtt
GAR1 siRNA II	CGUGUAGUCUUAUUAGGAGtt CUCCUAAUAAGACUACACGtt
GAR1 siRNA III	CGGAAUAGUGAAUUUUGCUtt AGCAAAAUUCACUAUUCCGtt
DKC1 siRNA I	GCUUUCUAGGGCCCUAGAAtt UUCUAGGGCCCUAGAAAGCtg
DKC1 siRNA II	GGACGGCAUUGAGGUCAAUtt AUUGACCUCAAUGCCGUCCtc
DKC1 siRNA III	CCUUUAUGUCCUAUUUGGAtt UCCAAAUAGGACAUAAAGGtc
DKC1 siRNA IV	GGACAACACACUAUACACCtt GGUGUAUAGUGUGUUGUCCtt
DKC1 siRNA V	GGACAGGUUUCAUUAAUCUtt AGAUUAAUGAAACCUGUCCtg
DKC1 siRNA VI	GGUUUCAUUAAUCUUGACAtt UGUCAAGAUUAAUGAAACCtg
Neg 1 - Ambion Negative control *	Unknown
Neg 2 - Invitrogen Negative control	Unknown
Neg 3 - Qiagen Negative control	Unknown
Neg 4 - GenePharma Negative control	UUC UCC GAA CGU GUC ACG UTT ACG UGA CAC GUU CGG AGA ATT

*Used in the transient knockdown experiments.

Appendix B – Primer Sequences

B.1. Miscellaneous primers

Name	5' to 3' Sequence
ABL 1	GGAGTGTTTCTCCAGACTGTTG
ABL 2	TTCAGCGGCCAGTAGCATCTGACTT
hU6 Forward	ACGCGTAGTGGAAAGACGCGCAGGCAAA
hU6 Reverse	GCGGCCGCTGTTTCGTCCTTTCCACAA
M13 Forward	TGTAAAACGACGGCCAGT
M13 Reverse	CAGGAAACAGCTATGAC

B.3. H/ACA RNA quantification primers (SYBR green)

Name	5' to 3' Sequence
U19 FWD	CGGTTGTCAGCTATCCAGG
U19 REV	CTCAGCTACAACACCAAAACAG
U64 FWD	CTCTCTCGGCTCTGCATAG
U64 REV	AAGAGAGGCCACAGTAAGG
U68 FWD	CCTAAACCCAAGAATCACTG
U68 REV	AAATTCACTTTGAGGGGCACG
U70 FWD	CAATTAAGCCGACTGAGTTCC
U70 REV	TCCCTTAGAGCAACCCATAC
U108 FWD	TTCAAATGGGCCTAACTCTGC
U108 REV	TGATGCAGGAAAGGCTTAAGAG
U3 FWD	CACGAGGAAGAGAGGTAGC
U3 REV	CCAATACGGAGAGAAGAACG
U97 FWD	TGCCCGATGATTATAAAAAGAC
U97 REV	AATCTTCGCTCACAGGACGCTG
ABL FWD	GATACGAAGGGAGGGTGTACCA
ABL REV	CTCGGCCAGGGTGTTGAA

B.3. TaqMan® primer/probe sequences and reaction concentrations

Primer/Probe	5' to 3' Sequence	Conc (nM)
ABL Forward	GATACGAAGGGAGGGTGTACCA	100
ABL Reverse	CTCGGCCAGGGTGTTGAA	100
ABL Probe	FAM-TGCTTCTGATGGCAAGCTCTACGTCTCCT-MGB	300
TERC Forward	GGTGGTGGCCATTTTTTGTC	300
TERC Reverse	CTAGAATGAACGGTGGAAGGC	300
TERC Probe	FAM-CGCGCTGTTTTTCTCGCTGACTTTC-MGB	200
NOP10 (Hs00430282_m1)	TaqMan® Gene Expression Assays (Applied Biosystems). Pre-mixed primer and probe set at 20x concentration. Individual sequence and concentration details undisclosed.	1x
NHP2 (Hs00750357_s1)		
NHP2.2 (Hs00950765_g1)*		
Dyskerin (Hs00154737_m1)		
GAR1 (Hs00255867_m1)		

*3' UTR specific - used to distinguish endogenous and exogenous *NHP2* expression.

B.4. *NOP10* shRNA oligonucleotide sequences

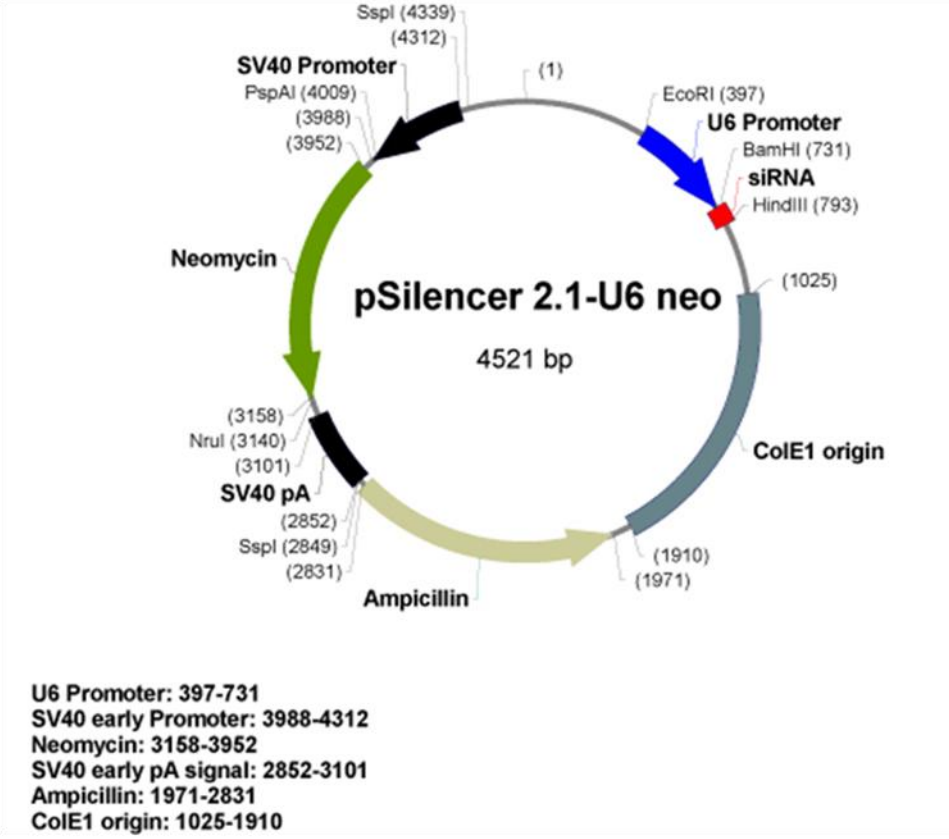
Name	5' to 3' Sequence
NOP10 shRNA 1	GGAACACATTGGTGTGGAATGTGTTCCCTTTATGGGCTCGAGGCGGCCGC
NOP10 shRNA 2	GAATTCGCGCCTTTTTTCTAGAGTCGACCCCATAAAGGGAACACATTGGT
NOP10 shRNA (+)	GATCCGCCCATAAAGGGAACACATTCCACACCAATGTGTTCCCTTTATGGGCTTTTTTTGGAAA
NOP10 shRNA (-)	AGCTTTTCCAAAAAAAGCCCATAAAGGGAACACATTGGTGTGGAATGTGTTCCCTTTATGGGCG

B.5. MMqPCR primers

Name	5' to 3' Sequence
telg	ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT
telc	TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTAACA
hbgu	CGGCGGCGGGCGGCGCGGGCTGGGCGGCTTCATCCACGTTCACCTTG
hbgd	GCCCGGCCCGCCGCGCCCGTCCCGCCGGAGGAGAAGTCTGCCGTT

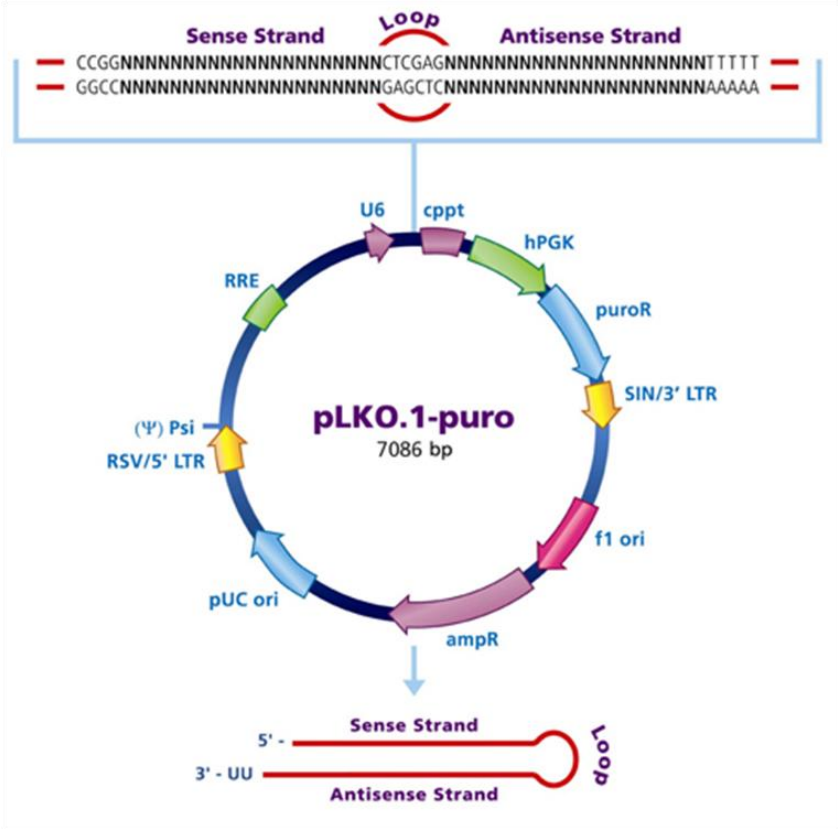
Appendix C – shRNA Expression Vectors

C.1. p*Silencer*™ 2.1-U6 neo



The p*Silencer*™ 2.1-U6 neo shRNA expression vector from Ambion, now part of Life Technologies Ltd, Paisley, UK. This map was taken from their website (http://www.lifetechnologies.com)

C.2. TRC1.5 MISSION® shRNA vector



cppt	Central polypurine tract
hPGK	Human phosphoglycerate kinase eukaryotic promoter
puroR	Puromycin resistance gene for mammalian selection
SIN/LTR	3' self inactivating long terminal repeat
f1 ori	f1 origin of replication
ampR	Ampicillin resistance gene for bacterial selection
pUC ori	pUC origin of replication
5' LTR	5' long terminal repeat
Psi	RNA packaging signal
RRE	Rev response element

The TRC1.5 MISSION shRNA lentiviral expression vector (based on pLKO.1-puro) from the Sigma-Aldrich Company Ltd, Dorset, UK. This vector map was taken from their website (http://www.sigmaaldrich.com).